76-18,778

KUHN, Steven Thomas, 1949-
MANY-SORTED MODAL LOGICS.

Stanford University, Ph.D., 1976
Philosophy

Xerox University Microfilms, Ann Arbor, Michigan 48106

MANY-SORTED MODAL LOGICS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHILOSOPHY
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Steven Thomas Kuhn
March 1976

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Kit Fine

(Principal Adviser)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ian Hacking

Approved for the University Committee on Graduate Studies:

W B Carnochan

Dean of Graduate Studies

ACKNOWLEDGMENTS

I am fortunate to have received help from a number of kind and clever people. Among these are Professors Dagfinn Follesdal, Ian Hacking, Stig Kanger, Julius Moravcsik, Krister Segerberg, Jonathan Stavi, and Patrick Suppes. Three people deserve special thanks: Professor Dov Gabbay, Professor Kit Fine, and Mr. Anthony Ungar. Professor Gabbay was my advisor and teacher at Stanford. Although he left Stanford before I began this dissertation, the influence of his teaching will be obvious to those who know his work. Professor Fine gave me a great deal of his time during the 1973-74 school year while under no official obligation to do so, and has continued to provide help and encouragement. The topic was Professor Fine's suggestion, and his advice along the way always proved sound. From Mr. Ungar I received the benefit of innumerable conversations on topics ranging from the nature of logic and philosophy to the fine points of English composition. His careful reading of the manuscript preserved me from the graduate student's perpetual fear -- that no one will look at what he has to say. Without it the dissertation would undoubtedly be worse.

Most of the writing of this dissertation was done at the Åbo Akademi in Finland and Uppsala University in Sweden while I was supported by a fellowship from the American-Scandanavian Foundation.

I am grateful to the people of both universities for their hospitality, and to the Foundation for making the trip possible.

Finally, my wife Betsey deserves credit for bearing so well the misfortune of having a husband who was writing a thesis at the same time she was.

TABLE OF CONTENTS

INTRODUCTION

This thesis is divided into two parts. Part I is intended to provide the background in modal logic needed for Part II. It contains a uniform treatment of modal logics (including classical logic as a special case). Consequence relations are emphasized, rather than the logics themselves, making it possible to dispense with many of the usual presuppositions about the formal language. The reader who already knows some modal logic may prefer to begin with Part II, and refer to Part I only as needed.

Part II deals with modal systems which are 'propositionally many-sorted'. The motivation for the study of these systems comes from tense logic. It has often been pointed out that the traditional tense logics do not provide a satisfactory analysis of English tenses. For example, they provide no means of distinguishing between sentences like 'John builds a house', which admit a progressive tense, and those like 'Two is a prime number', which don't. It seems reasonable to say that sentences of the first kind should be evaluated at intervals of time, those of the second, at instants.[1] The progressive tense of a sentence A is true at an instant if A is true at some

[1] As it turns out, this is a little too simple-minded. In the system of Chapter 3, sentences of the first kind are evaluated _at_ instants (of utterrance) _with reference to_ intervals (of occurrence). Sentences of the second kind are evaluated at instants with reference to instants.

interval containing that instant. If A is 'instant-evaluated', the progressive tense of A is ungrammatical. This explains why 'Two is a prime number' has no progressive tense and also why the operation of forming the progressive tense cannot be iterated. Systems like these in which different kinds of sentences are evaluated at different kinds of 'worlds' are exactly what we mean by 'propositionally many-sorted' systems. Such systems may have other applications as well (particularly in areas where the interpretation of iterated connectives has proved difficult).

In the first of the three chapters of Part II, a general theory of many-sorted systems is outlined. Their application to tenses is presented in the next chapter.

The final chapter deals with the interpretation of quantifiers in classical logic. Similarities between quantifiers and modal connectives have been noted many times. In this chapter we show that the work of quantifiers can, in fact, be done by propositional connectives within our many-sorted framework. What would normally be expressed by a formula of predicate calculus with n free variables can be expressed by a modal sentence evaluated at 'worlds' which are n-tuples of individuals. It turns out to be more convenient to base our systems on the variable-free equivalents of predicate logic than on predicate logic itself. We consider systems based on two such predicate logic-equivalents: Tarski's cylindric algebra and Quine's predicate functor logic. A third system is also introduced in order to bring out the resemblance between these systems and Segerberg's 'basic' two-dimensional modal logic, B. Each of the three systems is shown to be equivalent (in a sense defined in Part I) to predicate logic. This adds weight to Von Wright's view that quantification can be regarded as a modality. In the case of predicate functor logic, the modal approach has another virtue. It leads to the solution of a problem of Quine's, viz., axiomatizing predicate functor logic in such a way that its theorems correspond exactly to those of classical predicate logic.

CHAPTER I

LOGICAL FRAMEWORKS

A. Introduction

A logical framework consists of a language, a class of structures for the language and, for each structure, a class of possible interpretations of the language in the structure. The language can be thought of as an idealization of some part of a natural language. It is determined by a set of symbols and a set of rules by which the symbols can be concatenated to form sentences. The set of rules is such that it is always possible to determine in a finite number of steps whether or not a given string of symbols is a sentence. A structure for a language can be thought of as a (possibly unfaithful) representation of certain features of the world. An interpretation is a link between language and structure which makes it possible for sentences of the language to express information about the world. We can think of an interpretation as a procedure which enables us to determine whether sentences in the language are true or false "in the structure", ie. whether the sentences would be true or false if the world were as represented in the structure. To specify an interpretation usually requires two steps. First, we associate with each symbol of the language an object contained in the structure. The formal device which spells out this association is called a model. Then we outline a uniform procedure by which the truth of a sentence in a structure can be computed from the objects which are associated with its constituent symbols. Once this procedure is given all the information needed to determine the truth-in-s of sentences under i is provided by the model corresponding to i. Hence we usually talk about 'truth in a model' rather than truth in a structure under an interpretation.

From the preceding description it is apparent that a great deal of latitude is permitted in the choice of a logical framework. First, nothing was said about which part of language or which features of the world are to be captured. Furthermore, no guides were given for determining the class of structures. We can consider only structures which represent conditions similar to those which obtain in the real world, or we can allow structures which represent conditions so bizarre as to be unimaginable. Similarly, there were no restrictions placed on the set of interpretations. A symbol might have a wide range of possible values under different interpretation functions, or it might be assigned the same object by every possible interpretation. Finally, the manner in which truth values are computed from the interpretation functions is left completely open.

Despite their apparent diversity, the logical frameworks share an important property. Every logical framework determines a relation between sentences that corresponds to a notion of logical consequence and a set of sentences which corresponds to a notion of logical truth. A sentence A is a logical consequnce of a set of sentences Γ (with respect to a logical framework) if, in every structure, B is true under all interpretations under which all the members of Γ

are true (or equivalently, if, whenever every member of Γ is true in a model, so is B). A sentence is logically true or valid (with respect to a logical framework) if, in every structure, it is true under all interpretations (or equivalently, if it is true in all models).[1] Thus, if the aim of logic is to find criteria by which sentences can be classified as logically ture or not logically true, and by which sentences can be classified as logical consequences of sets of sentences or not, then this aim can be met by specifying a logical framework.

The criteria provided, however, may not be as useful as we would like. Since 'consequence' and 'validity' are defined in terms of <u>all</u> structures and interpretations or <u>all</u> models it may be impossible to determine in a finite number of steps whether a given sentence is valid or whether one sentence is a consequence of another. In fact, to check that a valid sentence is true in all models and to find a model which falsifies a non-valid sentence may both be tasks requiring unlimited time. For this reason we rely on the <u>axiomatic method</u>. Certain valid sentences are singled out as <u>axioms</u>. Then rules are given which enable us to generate other valid sentences from the axioms. We require that both the rules and axioms can be enumerated

---

[1]It might seem odd to define 'consequence' and 'validity' relative to a framework and to talk of <u>a</u> noticn of logical consequence instead of <u>the</u> notion of logical consequence. But a good case can be made for the thesis that there is no one notion of logical consequence, and that an explanation for some of the differences between the logics of, say, Leibnitz and Frege is that there are different logical frameworks underlying them. But these matters will not be discussed here.

in some effective way. Any sentence which can be obtained from the axioms by applying the rules a finite number of times is called a theorem. If a sentence is a theorem, there is an effective procedure for verifying this fact. First we list all the theorems that can be generated from axiom 1 by 1 or fewer applications of rule 1, then all the theorems that can be generated from axioms 1 and 2 by 2 or fewer applications of rules 1 and 2, and so on. It is clear that at each stage, we add only a finite number of theorems to the list. So we can check, at each stage, whether the theorem to be tested has appeared on the list. Since every theorem is generated from a finite number of axioms by a finite number of rules, we must eventually find our test theorem on the list.

This method enables us to recognize some logical truths, but, unless every logical truth is a theorem, it will not provide the kind of general verification procedure we wanted. Thus we are interested in finding axioms and rules which will make every logical truth a theorem. Similarly, we are interested in finding a way of generating all the pairs (Γ, B) such that B is a consequence of Γ from some finite set of these pairs. These tasks we call the tasks of axiomatizing a logic and axiomatizing a consequence relation, respectively.

Sometimes, we reverse this procedure. We start with the assumption that certain sentences are valid and that certain rules always generate valid sentences from other valid sentences (or that certain sentences are consequences of certain others and certain rules always generate consequence-related pairs from other consequence-related pairs). We then investigate classes of models which are compatible with these assumptions.

We have now talked in a very general way about logical frameworks, languages and models. In the remainder of this chapter we shall make these matters more precise for a number of special and well-known cases. We shall be concerned in particular with the problem of axiomatizing logics and consequence relations. We shall also investigate the relation between logical consequence and logical truth and the problem of determining when two consequence relations or logics are equivalent. Much of the material in this chapter will be used (or at least imitated) in the next chapter when a wider class of logical frameworks is introduced.

B. Languages

Definition 1.1a. A propositional language is a pair (S,C) where S is a set (the set of sentence letters) and C is a collection of pairwise disjoint (possibly empty) sets C_n for n a non-negative integer. The members of C_n are called n-ary connectives, and 0-ary connectives are also called sentence constants. No sentence letter is identical to any n-ary connective.

Definition 1.2a. If $L = (S,C)$ is a propositional language, the set of sentences of L is the smallest set X containing $S \cup C_0$ as a subset and closed under the rule:

(R1) If $A_1,\ldots,A_n$ are in X and $\Box$ is in C_n then the sequence $(\Box, A_1,\ldots,A_n)$ is in X.

Definition 1.1b. A predicate language is a triple (X,P,C) where X is a countable set (the set of individual variables), P is a collection of countable sets P^n for each non-negative integer n, C is as in Definition 1.1a except that, for each x in X, C_1 must contain the special connectives $\forall x$ and $\exists x$. These are called quantifiers. The members of P^n are called n-ary predicate letters. The 0-ary predicate letters are also called sentence letters. (If (X,P,C) is a predicate language then (P^0,C) is a propositional language with sentence letters in P^0.) The individual variables, predicate and sentence letters, and connectives are all distinct.

Definition 1.2b. If $L = (X,P,C)$ is a predicate language, then the atomic sentences of L are all finite sequences of the form (Q^0) or $(Q^n, x_1,\ldots,x_n)$ where Q^0 is in P^0, Q^n is in P^n, and $x_1,\ldots,x_n$ are in X. The sentences of L are the members of the smallest set which contains all the atomic sentences and is closed under (R1) above.

We represent sentences and atomic sentences by writing the names of their coordinates in order from left to right (with no commas or parentheses). We do not lose anything by this convention since none of the symbols we use to denote objects in the language can be got by writing two or more other such symbols side by side. An occurrence of a variable x in a sentence A is said to be bound if that occurrence is part of an occurrence of a sentence of the form $\forall x\, B$ or $\exists x\, B$ in A. Otherwise that occurrence is said to be free. We sometimes write '$A(x_1,\ldots,x_n)$' instead of 'A'

to indicate that $x_1,\ldots,x_n$ might have free occurrences in A. If we do this and later in the same discussion we introduce the notation '$A(y_1,\ldots,y_n)$', then '$A(y_1,\ldots,y_n)$' is understood to be the result of simultaneously substituting y_i for each free occurrence of x_i. We frequently identify a predicate or propositional language with the set of its sentences. By the <u>constituents of L</u> we mean the connectives, sentence letters, etc., which are the coordinates of sentences of L.

<u>Definition 1.3</u>. If A is a sentence of a propositional or predicate language then the <u>length of A</u> is the number of occurrences of connectives in A.

<u>Definition 1.4</u>. A propositional or predicate language is <u>Boolean</u> if it contains ' $-$ ' as a unary connective and ' $\rightarrow$ ', ' $\wedge$ ', ' $\vee$ ' as binary connectives. '$-$', '$\rightarrow$', '$\wedge$', and '$\vee$' are called <u>Boolean connectives</u>. We deviate from the conventions stated earlier by writing:

$$(A \wedge B), \quad (A \vee B) \quad \text{and} \quad (A \rightarrow B)$$

instead of

$$'\wedge'AB, \quad '\vee'AB, \quad \text{and} \quad '\rightarrow'AB,$$

respectively. Outermost parentheses in the representations of a sentence containing occurrences of Boolean connectives are normally omitted. Non-Boolean connectives which are not quantifiers are called <u>modal</u> connectives.

The word 'language' will be understood to mean predicate or propositional language throughout the remainder of this chapter.

C. Models

Definition 1.5. If W is a non-empty set, then for any non-negative integer n, an n-ary neighborhood relation on W is a subset of $W \times (2^W)^n$. If R is an n-ary neighborhood relation on w and $(w, U_1, \ldots, U_n) \in R$, we write: $Rw, U_1, \ldots, U_n$. (0-ary neighborhood relations on W are subsets of W.)

Definition 1.6. If W is a non-empty set then

$$\underline{\text{Neg}}\ \underline{W} = \{(w,U) : w \notin U\} ;$$

$$\underline{\text{Conj}}\ \underline{W} = \{(w,U,V) : w \in U \cap V\};$$

$$\underline{\text{Disj}}\ \underline{W} = \{(w,U,V) : w \in U \cup V\};$$

and
$$\underline{\text{Impl}}\ \underline{W} = \{(w,U,V) : w \in \overline{U} \cup V\} .$$

(Neg W is a unary neighborhood relation on W and Conj W, Disj W and Impl W are all binary neighborhood relations on W.)

Definition 1.7. If W is a non-empty set then a point-to-point relation on W is a subset of $W \times W$. If R is a point-to-point relation on W and $(u,v) \in R$, we write: uRv.

Definition 1.8a. If L is a propositional language, a modal model suitable for L is a 4-tuple $M = (W, 0, \overline{C}, V)$ such that

1) W is a non-empty set (the set of points of evaluation, or simply points).[2]
3) $\overline{C} = \{ \ddot{\Box}_M : \Box$ is a connective of $L\}$ where $\ddot{\Box}_M$ is either a point-to-point relation on W or a unary neighbor-

[2] We reserve the term 'worlds' for use in discussions of necessity and possibility.

hood relation on W if $\Box$ is a unary connective, and $\overline{\Box}_M$ is an n-ary neighborhood relation on W if $\Box$ is an n-ary connective of L.

2) 0 is a member of W.

4) V is a function (the <u>valuation</u>) which assigns to each sentence letter of L a subset of W.

<u>Definition 1.8b</u>. If L is a predicate language, a <u>modal model suitable for L</u> is a 6-tuple $M = (W, 0, D, \overline{P}, \overline{C}, a)$ such that clauses 1) and 2) of the previous definition hold and in addition

3) $\overline{C} = \{\overline{\Box}_M : \Box \text{ is a connective of } L\}$ where $\overline{\Box}_M$ is an n-ary neighborhood relation if $\Box$ is n-ary.[3]

4) D is a non-empty set whose members are called the <u>individuals</u> of M.

5) $\overline{P} = \{\overline{P}_M : P \text{ is a predicate letter of } L\}$ where, if P is n-ary, $\overline{P}_M$ is a function from W to subsets of D^n.

6) a is a function from the individual variables of L to D. a is called the <u>assignment function</u>. It is usually treated separately, but we take it to be part of the model in order to make our exposition conform to the description in the introduction. This treatment will also simplify later work.

<u>Definition 1.8c</u>. If L is a propositional language, an <u>algebraic</u>

---

[3]It is also possible to interpret connectives of a predicate language by point-to-point relations. These interpretations involve special complications, however, and will not be dealt with here. See [Gabbay, f], Chapters 2,5.

model suitable for L is a 4-tuple $M = (A, D, F, V)$ such that

1) A is a non-empty set (the set of truth values).
2) D is a non-empty subset of A (the designated truth values).
3) $F = \{F_\Box : \Box \text{ is a connective of } L\}$ where $F_\Box : A^n \to A$ if $\Box$ is n-ary.
4) V is a function (the valuation) from the sentence letters of L into A.

Definition 1.9a. Suppose L is a propositional language, $A \in L$, $M = (W, 0, \overline{C}, V)$ is a modal model suitable for L, and w is a point in M. Then A is true-at-w in M (written: $(M,w) \models A$) if one of the following holds:

1) A is a sentence letter of L and $w \in V(A)$.
2) $A = \Box B_1 \ldots B_n$ for some n-ary connective $\Box$ of L and either
   a) $n \geq 1$ and $\overline{\Box}_M\, w, \overline{B}_1, \ldots, \overline{B}_n$ where for $1 \leq j \leq n$, $\overline{B}_j = \{u \in W : (M,u) \models B_j\}$ or
   b) $n = 1$, $\overline{\Box}_M$ is a point-to-point relation and, for all u in W, $(M,u) \models B$ if $w\, \overline{\Box}_M\, u$ or
   c) $n = 0$ and $w \in \overline{\Box}_M$.

A is true in M (written: $M \models A$) if $(M,0) \models A$. Otherwise A is false in M (written: $M \not\models A$).

Definition 1.9b. Suppose L is a predicate language, $A \in L$, $M = (W, 0, D, \overline{P}, \overline{C}, a)$ is a modal model suitable for L, and $w \in W$. Then A is true-at-w in M (written: $(M,w) \models A$) if one of the following holds:

1) A is a 0-ary predicate letter and $w \in \overline{A}_M$.

2) For some $n \geq 1$ and some n-ary predicate letter Q,

$A = Q\, x_1 \ldots x_n$ and $(a(x_1), \ldots, a(x_n)) \in \overline{Q}_M(w)$.

3) For some $n \geq 0$ and some n-ary connective $\Box$ of L,

$A = \Box B_1 \ldots B_n$ and either clause 2a or clause 2b of the previous definition holds.

4) $A = \forall x\, B$, x is an individual variable and B is a sentence, and, for all modal models $M' = (W, 0, D, \overline{P}, Z\cdot, a')$ if $a'(y) = a(y)$ whenever $y \neq x$ then $(M', w) \models B$.

5) $A = \exists x\, B$, there is some function a' such that $a'(y) = a(y)$ whenever $y \neq x$, and $(M', w) \models B$ where $M' = (W, 0, D, P, \overline{C}, a')$.

<u>A is true in M</u> (written: $M \models A$) if $(M, 0) \models A$. Otherwise, <u>A is false in M</u> ($M \not\models A$).

<u>Definition 1.9c</u>. Suppose L is a propositional language, B is in L, and $M = (A, D, F, V)$ is an algebraic model suitable for L. Then the <u>value of B in M</u> (written: $V^*(B)$) is defined as follows:

1) If B is a sentence letter then $V^*(B) = V(B)$.

2) If $B = \Box C_1 \ldots C_n$ for some n-ary connective $\Box$, then $V^*(B) = F_\Box (V^*(C_1), \ldots, V^*(C_n))$.

<u>B is true in M</u> ('$M \models B$') if $V^*(B) \in D$. Otherwise, <u>B is false in M</u> ($M \not\models B$).

In a modal propositional model the world is pictured as a collection of different situations which are related to each other in various ways. (The actual situation is represented by 0.) For example, in the literature models have been considered in which two situations are related by the fact that the first has obtained or will obtain before the second, by the fact that the second could possibly obtain given that the first actually obtains, and by the fact that the first is more similar to the situation which actually obtains than the second. Each sentence is either true or false relative to each situation, and whether a complex sentence is true relative to a situation depends on whether its components are true at related situations. A sentence is true simpliciter if it is true relative to the situation which actually obtains. In a modal predicate model the picture is refined by the addition of a stock of objects (to which the individual variables refer) and relations among the object (to which the predicate letters refer). The objects are taken to be present in every situation, though the relations among them may change. It is not so clear how an algebraic model represents the world. We consider these models mainly as a tool for investigating the modal ones. We shall see later that modal models can always be replaced by algebraic ones.

Notation. If M is a modal model suitable for the propositional language L we often write 'W_M', 'O_M', '$\overline{C}_M$', and 'V_M' for the first, second, third, and fourth coordinates of M, respectively. When confusion is unlikely the subscripts are dropped. If $\boxed{1}$, $\boxed{2}$, ... are all the connectives of L then we sometimes write

$$(W_M, O_M, \overline{\boxed{1}}_M, \overline{\boxed{2}}_M, \ldots, V_M) \text{ for } M.$$

Similar conventions apply to modal models suitable for predicate languages and to algebraic models. We shall not need to talk about algebraic models very often, so unless it is explicitly stated otherwise 'model' is used as an abbreviation for 'modal model'. If M is a model suitable for L, $w \in W_M$, and $\Gamma \subseteq L$, $\Gamma \neq \phi$, then we write $(M,w) \models \Gamma$ if, for all A in Γ, $(M,w) \models A$. Similarly, $M \models \Gamma$ if, for all A in Γ, $(M,0) \models A$. It is also convenient to stipulate that $M \models \phi$.

Definition 1.10. If M is a model suitable for L, then M is trivial if either, for all sentences A of L, $M \models A$ or, for all sentences A, $M \not\models A$. Trivial models do not discriminate at all among the sentences of the language. It will sometimes be convenient to exclude them.

Definition 1.11a. If L is a propositional (predicate) language and M is a model suitable for L, then M is partially classical if the following hold:

1) If '$-$' is a unary connective of L, then $\bar{-}_M = \text{Neg } W_M$.
2) If '$\rightarrow$' is a binary connective of L, then $\bar{\rightarrow}_M = \text{Impl } W_M$.
3) If '$\wedge$' is a binary connective of L, then $\bar{\wedge}_M = \text{Conj } W_M$.
4) If '$\vee$' is a binary connective of L, then $\bar{\vee}_M = \text{Disj } W_M$.

Furthermore, if L contains '$\wedge$', '$\vee$', and '$\rightarrow$' as binary connectives then M is deductive, and if L contains all the Boolean connectives then M is classical.

Definition 1.11b. If L is a propositional language and M is an algebraic model suitable for M, then M is classical if L is Boolean, for some t, $D_M = \{t\}$ and $(A, F_\wedge, F_\vee, F_-, t, F_-(t))$ is a Boolean algebra, i.e., $F_\wedge$ and $F_\vee$ are commutative and associative, $F_\wedge$ distributes over $F_\vee$, $F_\vee$ distributes over $F_\wedge$, and, for all a, b in A, $F_\wedge(F_\vee(a,b),b) = b$, $F_\vee(F_\wedge(a,b),b) = b$, $F_\wedge(a,F_-(a)) = t$, and $F_-(a, F(a)) = F_-(t)$.

From definitions 1.5, 1.8a, 1.8b, 1.9a, 1.9b, we know the following.

<u>Lemma 1.1.</u> If L is a language, M is a partially classical model suitable for L, and $w \in W_M$, then for all A, B in L:

1) If '-' is a unary connective of L, then $(M,w) \models -A$ iff $(M,w) \not\models A$.

2) If '→' is a binary connective of L, then $(M,w) \models A \rightarrow B$ iff $(M,w) \not\models A$ or $(M,w) \models B$.

3) If '∧' is a binary connective of L, then $(M,w) \models A \wedge B$ iff $(M,w) \models A$ and $(M,w) \models B$.

4) If '∨' is a binary connective of L, then $(M,w) \models A \vee B$ iff $(M,w) \models A$ or $(M,w) \models B$.

If L is a predicate language, M' is a model suitable for L, $W_{M'} = W_M$, $O_{M'} = O_M$, $D_{M'} = D_M$, $\overline{P}_{M'} = \overline{P}_M$, $\overline{C}_{M'} = \overline{C}_M$, then the following also hold:

5) If $x_1,\ldots,x_n$ are among the free variables of $A(y_1,\ldots,y_n)$, $y_1,\ldots,y_n$ have no occurrences in $A(x_1,\ldots,x_n)$ and for $1 \leq i \leq n$ $a_M(x_i) = a_{M'}(y_i)$ then $(M,w) \models A(x_1,\ldots,x_n)$ iff $(M',w) \models A(y_1,\ldots,y_n)$.

6) If $(M,w) \models \exists x\, A(x)$ and y has no occurrences in $\exists x\, A(x)$ then $(M,w) \models \exists y\, A(y)$.

7) If x does not occur free in A then $(M,w) \models \exists x\, A$ iff $(M,w) \models A$ iff $(M,w) \models \forall x\, A$.

Proof. 1-4 follow immediately from the definitions quoted. 5 is proved by induction on the length of A. 6 and 7 are consequences of 5.

In view of 1-4 of the preceding theorem, we do not need to know anything about W_M or 0_M to determine the truth in M of sentences which only contain Boolean connectives and quantifiers. For this reason if every connective of L is Boolean, or a quantifier, then we usually drop the first two coordinates in exhibiting a partially classical model suitable for L. Further savings are obtained by dropping $\bar{\wedge}$, $\bar{\vee}$, $\bar{\rightarrow}$, $\bar{-}$, from a partially-classical model M suitable for L (regardless of what other connectives are in L). Thus a classical model suitable for a predicate language is written '(D,P,a)'. The $\bar{P}$'s in P can then be regarded simply as subsets of D^n.

Note: The analogs of (1)-(4) above also hold for algebraic models suitable for L. This follows from some elementary facts about Boolean algebras. (See, for example, [Rasiowa, Sikorski, 1963].)

We conclude this section with two lemmas to be used later.

Definition 1.12. If M is a model suitable for L and A is a sentence of L, then the truth set of A in M (written $|A|_M$) is the set of all w in W_M such that $(M,w) \models A$. Any subset of W_M which is the truth set of a sentence of L is said to be <u>definable in M</u>.

<u>Lemma 1.2</u>. Let M and M' be two models for L which are alike except that $\bar{\Box}_M \neq \bar{\Box}_{M'}$ for some n-ary connective $\Box$ of L. Suppose for all sets $U_1,\ldots,U_n$ which are definable in M and all $w \in W_M$ $\bar{\Box}_M\, w, U_1,\ldots,U_n$ iff $\bar{\Box}_{M'}\, w, U_1,\ldots,U_n$. Then, for all A in L, $M \models A$ iff $M' \models A$.

<u>Proof</u>. Straightforward induction on the length of A.

<u>Lemma 1.3</u>. If M is a modal model suitable for L then there is an algebraic model $\mathfrak{A}(M)$ suitable for L such that, for all B in L $\mathfrak{A}(M) \models B$ iff $M \models B$.

<u>Proof</u>. Let $M = (W, 0, \bar{C}, V)$. If for all A in L, $(M,0) \not\models A$, then the algebraic model $(\{0,1\}, \{1\}, F, U)$ will do the job where U and each $F_\Box$ has range $\{0\}$. So we can assume that for some A $(M,0) \models A$. Let $A = \{|C|_M : C \in L\}$. Let $D = \{X \in A : 0 \in X\}$. (This is non-empty by the previous remark.) For all sentence letters C in L let $U(C) = |C|_M$. Finally, if $\Box$ is an n-ary connective of L, let $F_\Box(|C_1|, \ldots, |C_n|_M) = |\Box C_1 \ldots C_n|_M$. ($F_\Box$ is well defined because if $|B_i| = |C_i|$ for $1 \leq i \leq n$, then $\bar{\Box}\, w, |C_1|,\ldots,|C_n|$ iff $\bar{\Box}\, w, |D_1|,\ldots,|D_n|$ and hence $|\Box C_1 \ldots C_n| = |\Box D_1 \ldots D_n|$. Now let $\mathfrak{A}(M) = (A, D, G, U)$ where $G = \{G_\Box : \Box$ is a connective of $L\}$. The lemma follows by straightforward induction on the length of B.

D. Truth and Consequences

Definition 1.13. Suppose $\mathcal{M}$ is a class of models suitable for L and $B \in L$. Then B is logically true with respect to $\mathcal{M}$ (written: $\models_{\mathcal{M}} B$) if, for all M in $\mathcal{M}$, $M \models B$.

When we say that B is logically true we mean at least that B must be true no matter what conditions happen to prevail in the world. If our definition is to be faithful to this intuition we must insist that the classes of models we consider be very wide. We list below some reasonable looking conditions.

Definition 1.14. Let $\mathcal{M}$ be a class of models suitable for a language L.

a) $\mathcal{M}$ is normal if it is closed under the operation of shifting the designated point of any of its members.

b) If L is propositional, $\mathcal{M}$ is valuation-unrestricted if it is closed under the operation of changing a member's valuation function.

c) If L is predicate, $\mathcal{M}$ is assignment-unrestricted if it is closed under the operation of changing a member's assignment function.

Notation. If $\mathcal{M}$ is normal, $M \in \mathcal{M}$ and $w \in W_M$, we write M^w for the model obtained from M by replacing 0_M by w.

For the purpose of axiomatizing a consequence relation it turns out to be convenient to stretch the notion of consequence slightly, and call a <u>set</u> of sentences Δ a consequence of Γ if every condition which makes everything in Γ true makes something in Δ true.[4]

<u>Definition 1.15</u>. Suppose $\mathcal{M}$ is a class of models suitable for L and $\Gamma \subseteq L$, $\Delta \subseteq L$. Then <u>Δ is a logical consequence of Γ with respect to $\mathcal{M}$</u> (written $\Gamma \models_{\mathcal{M}} \Delta$) if for all M in $\mathcal{M}$, $M \models \Gamma$ implies there is some B in Δ such that $M \models B$. $\underline{\models_{\mathcal{M}}} = \{(\Gamma,\Delta) : \Gamma \models_{\mathcal{M}} \Delta\}$.

We close this section with a simple fact that is important enough to be listed as a lemma.

<u>Lemma 1.4</u>. If $\mathcal{M} \subseteq \mathcal{M}'$, then $\models_{\mathcal{M}'} \subseteq \models_{\mathcal{M}}$.

Henceforth we take L to be a language and $\mathcal{M}$ to be a class of models suitable for L.

E. Consequence Relations

In the next two sections we list some definitions which will be needed to solve the problem of axiomatizing logics and consequence relations determined by some natural classes of models. In this section, we deal with consequence relations.

To axiomatize the notion of logical consequence with respect to $\mathcal{M}$ we must first pick out a set of pairs (Γ,Δ) such that $\Gamma \models_{\mathcal{M}} \Delta$

---

[4] This kind of definition was recommended in [Scott, 1971].

holds, and then give some rules which enable us to obtain, from a sequence of pairs (Γ_i,Δ_i) such that $\Gamma_i \models_M \Delta_i$, a pair (Γ,Δ) such that $\Gamma \models_M \Delta$. But without knowing anything about our class of models, except that each model somehow assigns truth or falsity to every sentence, we can already list some of these axioms and rules.

Lemma 1.5. If $\vdash\, = \,\models_M$ then, for all $\Gamma, \Delta, \Psi, \Theta \subseteq L$:

(1) If $\Gamma \cap \Delta \neq \phi$ then $\Gamma \models \Delta$.

(2) If $\Gamma \vdash \Delta$ then $\Gamma \cup \Psi \vdash \Delta \cup \Theta$.

(3) If $\Gamma \vdash \Delta \cup \{A\}$ and $\{A\} \cup \Psi \vdash \Theta$ then $\Gamma \cup \Psi \vdash \Delta \cup \Theta$.

(We occasionally refer to properties (2) and (3) as 'expansion' and 'cut', respectively.) These properties were not chosen at random.) We will use them in all our axiomatizations of $\models_M$ for particular M. Furthermore we shall show later than any relation which satisfies (1)-(3) and one other condition can be regarded as $\models_M$ for some M. These considerations motivate the following definition.

Definition 1.16. A consequence relation on L is a binary relation $\vdash$ between sets of sentences of L such that for all $\Gamma \subseteq L$ and all $\Delta \subseteq L$ (1)-(3) of Lemma 1.2 hold.

Corollary. $\models_M$ is a consequence relation on L. Let us call it the consequence relation determined by M.

Notice that intersections of consequence relations are consequence relations. Hence if P is a property preserved under intersections which holds of some consequence relations, then it makes sense to talk about the smallest consequence relation satisfying P. Since the set of all pairs of sets of sentences of L is a consequence relation on L, it always makes sense to talk about the smallest consequence relation which contains a certain set of pairs and which is closed under certain rules.

If $\vdash$ is a consequence relation and $\Gamma \vdash \Delta$ then call (Γ,Δ) virtually finite (with respect to $\vdash$) if there are finite or empty sets $\Gamma' \subseteq \Gamma$, $\Delta' \subseteq \Delta$ such that $\Gamma' \vdash \Delta'$. Notice that all the pairs which can be shown to be $\vdash$-related by clause 1 are virtually finite and clauses ii and iii generate only virtually finite pairs if they are applied to virtually finite pairs. Nearly all the rules which we will consider later possess this property of "preserving virtual finiteness". Therefore the consequence relations we consider will all be finitary, i.e., we will have $\Gamma \vdash \Delta$ if and only if, for some finite or empty sets $\Gamma' \subseteq \Gamma$, $\Delta' \subseteq \Delta$, $\Gamma' \vdash \Delta'$. This property will prove important. It is needed for the reduction of consequence relations to logics (section F), and for the Lindenbaum lemma used to prove completeness (section G) and for many of the results in section H.

If all the models in M are non-trivial, then we can add to Lemma 1.5 the clause

(4) If $\Gamma \vdash \{A\}$ for all A in L, then $\Gamma \vdash \phi$.

If $\{A\} \vdash \Delta$ for all A in L, then $\phi \vdash \Delta$.

We call a consequence relation $\vdash$ which satisfies (4) regular. The rules expressed by (4) do not preserve virtual finiteness, for this reason we try to avoid including them in our axiomatizations, even though all of the consequence relations we shall consider turn out to be regular. Regularity is also needed for the reduction of consequence relations to logics.

We now define several special kinds of consequence relations and classify some connectives according to their behavior under consequence. When confusion is unlikely we write A for $\{A\}$; Γ,Δ for $\Gamma \cup \Delta$; and $\Box\Gamma$ for $\{\Box A : A \in \Gamma\}$. Also $\Gamma \dashv\vdash \Delta$ is short for $\Gamma \vdash \Delta$ and $\Delta \vdash \Gamma$.

<u>Definition 1.17</u>. Let L be a language and $\vdash$ be a consequence relation on L. Then $\vdash$ is <u>partially classical</u> if the following all hold.

(CC) If '$\wedge$' is a binary connective of L, then $\Gamma,A,B \vdash \Delta$ iff $\Gamma,A \wedge B \vdash \Delta$.

(CD) If '$\vee$' is a binary connective of L, then $\Gamma \vdash A,B,\Delta$ iff $\Gamma \vdash A \vee B,\Delta$.

(CI) If '$\rightarrow$' is a binary connective of L, then $\Gamma,A \vdash B,\Delta$ iff $\Gamma \vdash A \rightarrow B,\Delta$.

(CN) If '$-$' is a unary connective of L, then $\Gamma,A \vdash \Delta$ iff $\Gamma \vdash -A,\Delta$ and $\Gamma \vdash B,\Delta$ iff $\Gamma,-B \vdash \Delta$.

If L contains '$\wedge$', '$\vee$', '$\rightarrow$', or '$-$' and $\vdash$ is partially classical we say $\vdash$ has <u>classical conjunction</u>, <u>classical disjunction</u>, <u>classical implication</u> or <u>classical negation</u>, respectively. If $\vdash$ has classical conjunction, classical disjunction, and classical implication, we say $\vdash$ is <u>deductive</u>. If, in addition, L has classical negation we say $\vdash$ is <u>classical</u>.

The definition of deductive consequence relations was written in order to facilitate later developments. In particular, we can see at a glance that if $\vdash$ is deductive, then $\Gamma, A_1,\ldots,A_m \vdash B_1,\ldots,B_n,\Delta$ iff $\Gamma \vdash (A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m)\ldots) \to (B_1 \vee \ldots \vee (B_{n-1} \vee B_n)\ldots),\Delta$. It should be pointed out, however, that we could have written it in a more familiar-looking way.

<u>Lemma 1.6</u>. If L and $\vdash$ are as above, then $\vdash$ is deductive if and only if it satisfies:

(1') If $\Gamma,A,B \vdash \Delta$ then $\Gamma,A \wedge B \vdash \Delta$;

(1") If $\Gamma \vdash \Delta,A$ and $\Gamma \vdash \Delta,B$ then $\Gamma \vdash \Delta,A \wedge B$;

(2') If $\Gamma \vdash A,B,\Delta$ then $\Gamma \vdash A \vee B,\Delta$;

(2") If $\Gamma,A \vdash \Delta$ and $\Gamma,B \vdash \Delta$ then $\Gamma,A \vee B \vdash \Delta$;

(3') If $\Gamma,A \vdash B,\Delta$ then $\Gamma \vdash A \to B,\Delta$;

(3") If $\Gamma \vdash A,\Delta$ and $\Gamma,B \vdash \Delta$ then $\Gamma,A \to B \vdash \Delta$.

<u>Proof</u>. (a) Suppose $\vdash$ is deductive. Then (1'), (2'), (3') are satisfied and we need only check (1"), (2"), and (3"). We check (1") as an example. Applying (1) to $A \wedge B \vdash A \wedge B$ gives $A,B \vdash A \wedge B$. If $\Gamma \vdash \Delta,A$ then by cut $\Gamma,B \vdash A \wedge B$. If, in addition $\Gamma \vdash \Delta,B$ then by cut $\Gamma \vdash A \wedge B,\Delta$.

(b) Suppose (1'), (2'), (3'), (1"), (2"), (3") all hold. We must check (1), (2), (3) for right to left only. We do (1) as an example. Suppose $\Gamma,A \wedge B \vdash \Delta$. Since $\Gamma,A,B \vdash A$ and $\Gamma,A,B \vdash B$ we can apply (1') to get $\Gamma,A,B \vdash A \wedge B$. By cut, therefore, $\Gamma,A,B \vdash \Delta$.

Definition 1.18. If $\vdash$ is a consequence relation on L and $\Box$ is an n-ary connective of L, then

(a) $\Box$ is Montague with respect to $\vdash$ if $\vdash$ is closed under the rule "If $A_1 \dashv\vdash B_1$, and ... and $A_n \dashv\vdash B_n$, then $\Box A_1 \ldots A_n \dashv\vdash \Box B_1 \ldots B_n$".

(b) $\Box$ is normal-Kripke with respect to $\vdash$ if $n = 1$ and $\vdash$ is closed under "If $\Gamma \vdash A$, then $\Box\Gamma \vdash \Box A$".

(c) $\Box$ is normal-K4 with respect ot $\vdash$ if it is normal-Kripke with respect to $\vdash$ and, in addition, $\vdash$ is closed under "If $\Box\Gamma \vdash A$ then $\Box\Gamma \vdash \Box A$".

(d) $\Box$ is normal-S4 with respect to $\vdash$ if it is normal-K4 with respect to $\vdash$ and, in addition, $\vdash$ is closed under "If $A, \Gamma \vdash \Delta$ then $\Box A, \Gamma \vdash \Delta$".

(e) $\Box$ is normal-S5 with respect to $\vdash$ if it is normal-S4 with respect to $\vdash$ and, in addition, $\vdash$ is closed under "If $\Box\Gamma \vdash \Box\Delta, A$ then $\Box\Gamma \vdash \Box\Delta, \Box A$".

Definition 1.18f. If $\vdash$ is as above and $\Box$ and $\boxminus$ are unary connectives of L, then $(\Box, \boxminus)$ is a pair of tense connectives with respect to $\vdash$ if $\Box$ and $\boxminus$ are normal-S4 with respect to $\vdash$ and $\vdash$ is closed under the rules "If $\Box\Gamma \vdash \boxminus\Delta, A$ then $\Box\Gamma \vdash \boxminus\Delta, \Box A$ and "if $\boxminus\Gamma \vdash \Box\Delta, A$, then $\boxminus\Gamma \vdash \Box\Delta, \boxminus A$."

The motivation behind these definitions will not be clear until the completeness theorem has been proved. It turns out, however, that a connective $\Box$ which is normal-S4 or normal-S5 with respect to $\vdash$ can be interpreted as expressing some kind of necessity. We can think of the points in models which determine $\vdash$ as "possible worlds". The relations which interpret such connectives hold between two possible worlds if the second could have been the case if the first actually were the case, or more briefly if the second is possible relative to the first. $\Box A$ (read "necessarily A" in this case) is true at a world w if A is true at all worlds possible relative to w. In the case of S4 we require only that all worlds be possible relative to themselves, and that u's being possible relative to v and v's being possible relative to w entail u's being possible relative to w. In the case of S5-connectives we add to this the requirement that u is possible relative to v whenever v is possible relative to u. As we might expect from their rather simple definitions, S4 and S5 connectives may have other interpretations as well. It is the necessity interpretation, however, that has received the greatest amount of attention.

The normal-K4 and normal-Kripke connectives can also be interpreted by point-to-point relations. For these connectives, however, the relations need not be reflexive, so an arbitrary x normal-Kripke or normal-S4 connective cannot be thought of as expressing necessity. If a connective is merely normal-Kripke, it can be interpreted by a relation which lacks transitivity as well. Thus when we say a connective is normal-Kripke or normal-K4, we do not place a very great restriction on how it may be interpreted.

If $(\Box, \boxminus)$ is a pair of tense connectives with respect to $\vdash$ then we can think of these connectives as expressing

"It is and always will be the case that ..." and "It is and always was the case that ... ."[5] (It doesn't matter which phrase we match with which connective.) The points can be thought of as the world at different moments in its history (where "history" is construed broadly enough to include both past and future), and the relations which interpret $\Box$ and $\boxminus$, as the relations "earlier than or contemporaneous with" and "later than or contemporaneous with."

Finally, the Montague connectives turn out to be just the ones that can be handled within our framework (i.e., interpreted by a neighborhood relation). This is an extremely weak condition and it might be supposed that any connective whose logical properties warrant investigation would be Montague. It has been pointed out, however, that a connective suitable to represent the phrase "It is believed that ..." would not.

Definition 1.19. Let $\vdash$ and L be as above.

a) $\vdash$ is Montague if, for every connective $\Box$ of L, $\Box$ is Montague with respect to $\vdash$.

b) $\vdash$ is normal-Kripke if, for every non-Boolean connective $\Box$ of L, $\Box$ is normal-Kripke with respect to $\vdash$.

c) If L is propositional, $\vdash$ is substitution-closed if

---

[5]It should be noted that in what is usually called a tense logic one connective, G, is taken to mean "It will always be the case that..." and another connective H is taken to mean "It always was the case that..." Unlike the usual definitions, our definition of tense connectives has the advantage that it can be formulated without supposing anything about the behavior of the other connectives of the language. If $\vdash$ is classical $\Box A$ can be expressed by the conventional $GA \vee A$.

$\Gamma \vdash \Delta$ implies $\Gamma' \vdash \Delta'$ where Γ' and Δ' are the result of substituting a sentence A for every occurrence of some sentence letter p in Γ and Δ.

Definition 1.20. If $\vdash$ is a consequence relation on a predicate language then we say $\vdash$ has classical quantifiers if all the following hold whenever y does not occur free in $A(x)$:

a) If $\Gamma \vdash \Delta, A(x)$ then $\Gamma \vdash \Delta, \exists y A(y)$.

b) If $\Gamma, A(x) \vdash \Delta$ and x does not occur free in Γ or Δ, then $\Gamma, \exists y A(y) \vdash \Delta$.

c) If $\Gamma \vdash \Delta, A(x)$ and x does not occur free in Γ or Δ, then $\Gamma \vdash \Delta, \forall y A(y)$.

d) If $\Gamma, A(x) \vdash \Delta$ then $\Gamma, \forall y A(y) \vdash \Delta$.

HENCEFORTH WE ALWAYS TAKE $\vdash$, $\vdash'$ TO BE MONTAGUE CONSEQUENCE RELATIONS ON THE LANGUAGES L, L'. IF L, L' ARE PREDICATE LANGUAGES, WE ALSO ASSUME $\vdash$, $\vdash'$ HAVE CLASSICAL QUANTIFIERS.[6]

F. Logics

Writers of articles and textbooks on logic usually say very little about consequence relations and a great deal about "logics." A consequence relation was defined as a collection of pairs (Γ_i, Δ_i) such that (with respect to some logical framework) Δ_i

---

[6]Although we consider only languages with classical quantifiers, more general definitions are possible. See, for example [Gabbay, 1974].

is a logical consequence of Γ_i. Similarly, a <u>logic</u> is intended to be a collection Γ of sentences such that (with respect to some framework) Γ is the set of all logical truths. It is, of course, more convenient to deal with sentences than with pairs of sets of sentences, so we might think the logic writers should be forgiven for slighting the consequence relation. It turns out, however, that if a logic is identified with a set of sentneces, then there are many important notions which can't be properly formulated in terms of logics alone. For example, there does not seem to be any general way to define what it means for a sentence A of L, or a set Γ of sentences of L to be consistent with respect to a logic $\mathcal{L}$.[7] Sometimes it is said that A is $\mathcal{L}$-consistent iff -A is not a member of $\mathcal{L}$ (or iff there is some sentence B such that $A \rightarrow B$ is not a member of $\mathcal{L}$), and that Γ is $\mathcal{L}$-consistent iff, for every finite subset $\{A_1,\ldots,A_n\}$ of Γ, the sentence $(A_1\&(\ldots\&A_n)\ldots))$ is $\mathcal{L}$-consistent. But this account depends on $\mathcal{L}$'s having particular connectives, interpreted in a particular way.

A popular way around this difficulty is to use the term "logic" to refer not to a mere set of sentences, but to that set together with a set of "rules of derivation" which state that certain sentences are "immediately derivable" from certain sets

---

[7]The same goes for properties like 'strong completeness of $\mathcal{L}$' whose definitions depend on the notion of $\mathcal{L}$-consistency.

of sentences.[8] We can then say that a set of sentences is L-consistent if there is some sentence not derivable from $\Gamma \cup L$ by a finite number of applications of the rules of derivation of L.[9] On this approach, however, logics lose the advantage of being as simple and easy to deal with as sets of sentences. Furthermore, we don't really want to distinguish two logics just because they are associated with different rules of derivation. The only property of these rules we are really interested in (and the only one used in the definition above) is "Which sentences are derivable from which sets?" (We don't care, for example, how many rules there are, or whether derivations tend to be long or short.) But to know this property is just to know a consequence relation. For let $\Gamma \vdash \Delta$ iff either some sentence in Δ is derivable from Γ or Δ is empty and every sentence is derivable from Γ. Whatever the individual rules of derivation are like, we can be sure: that A is derivable from a set Γ containing A, that B's derivability from Γ and C's derivability from a set Γ' containing Γ entail C's

---

[8]Rules of derivation should not be confused with the rules discussed earlier, i.e., the rules which are used to generate the set L. Axiomatizations of classical predicate logic (Pred) often include the following two rules:

(Gen) if $A \in \text{Pred}$ then $\forall x\, A \in \text{Pred}$;

(MP) if $A \in \text{Pred}$ and $A \rightarrow B \in \text{Pred}$ then $B \in \text{Pred}$.

The second of these corresponds to a familiar <u>rule of derivation</u>, namely "from A and A $\rightarrow$ B, derive B." The first, however, does not.

[9]It should be noted that this definition and the one given before do not always agree. For example, if L contains all the classical propositional tautologies, but no rules of derivation, then p&-p is L-consistent according to this definition, but not according to the previous one.

derivability from $\Gamma \cup (\Gamma' - \{B\})$, and that B's derivability from Γ entails B's derivability from supersets of Γ. From these facts it is easy to establish that $\vdash$ is a consequence relation. Hence to say that B is deriable from Γ is just to say that (Γ, B) is a member of a consequence relation.[10]

We shall therefore stick to the idea that a logic is a set of sentences, admitting that it will often be necessary to bring in consequence relations. Under some conditions we can replace talk about consequence relations by talk about logics. In the remainder of this section we shall state some conditions under which this reduction can be carried out. We shall also verify that our classification of connectives according to their behavior under consequence relations corresponds to a well known classification according to their behavior in logics.

Definition 1.21. If $\vdash$ is a consequence relation on a language L, then the logic determined by $\vdash$ ($\underline{L}(\vdash)$) is the set of all sentences A of L such that $\phi \vdash A$.

Definition 1.22. A set of sentences Γ is a deductive (classical, regular) logic if it is the logic determined by a deductive (classical,regular) consequence relation.

---

[10] This argument does not show the need to consider consequence relations with plural sets on the right. The arguments against restricting the right hand side to singleton sets are apparently pragmatic. See [Scott, 1975].

Definition 1.23. If L is a deductive, regular logic, the finitary consequence relation determined by L (written: $\vdash_L$) is the set of all pairs Γ,Δ such that one of the following hold: (i) There is a finite set of formulas $\{A_1,\ldots,A_m\} \subseteq \Gamma$ and a finite set of formulas $\{B_1,\ldots,B_n\} \subseteq \Delta$ such that $(A_1 \wedge \ldots \wedge (A_{n-1} \wedge A_n)\ldots) \rightarrow (B_1 \vee \ldots \vee (B_{n-1} \vee B_n))$ is in L; or (ii) there is a finite set of formulas $\{B_1,\ldots,B_n\} \subseteq \Delta$ such that $(B_1 \vee \ldots \vee (B_{n-1} \vee B_n)\ldots)$ is in L; or (iii) there is a finite set of formulas $\{A_1,\ldots,A_m\} \subseteq \Gamma$ such that for every formula B $(A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m)\ldots) \rightarrow B$ is in L.

Lemma 1.7. (a) If L is a deductive regular logic then $\vdash_L$ is a finitary, deductive, regular consequence relation which determines L. (b) If $\vdash$ and $\vdash'$ are finitary, deductive, regular consequence relations and $L(\vdash) = L(\vdash')$ then $\vdash = \vdash'$.

Proof of (b). Suppose $\vdash$ and $\vdash'$ are finitary, deductive, regular consequence relations such that $L(\vdash) = L(\vdash')$ and that $\Gamma \vdash \Delta$. Since $\vdash$ is finitary there must be finite or empty sets $\Gamma' \subseteq \Gamma$, $\Delta' \subseteq \Delta$ such that $\Gamma' \vdash \Delta'$. If Γ' and Δ' are both non-empty then by deductiveness $\phi \vdash (A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m)\ldots) \rightarrow (B_1 \vee \ldots \vee (B_{n-1} \vee B_n)\ldots)$ where the A_i's run through Γ' and the B_j's run through Δ'. Since $\vdash'$ also determines L, we know that $\phi \vdash' (A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m)\ldots) \rightarrow (B_1 \vee \ldots \vee (B_{n-1} \vee B_n)\ldots)$. And, since $\vdash'$ is also deductive, $\Gamma' \vdash' \Delta'$. Hence $\Gamma \vdash' \Delta$. If Γ' is empty then we have $\phi \vdash \Delta'$. Since $\vdash$ is deductive, then, $\phi \vdash (B_1 \vee \ldots \vee (B_{n-1} \vee B_n))$ and therefore, $\phi \vdash' (B_1 \vee \ldots \vee (B_{n-1} \vee B_n))$.

Hence $\Gamma \vdash' \Delta$. If Δ' is empty, then we proceed similarly to get $(A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m) \ldots) \vdash \phi$. Therefore, for any formula B, $(A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m) \ldots) \vdash B$ so, since $\vdash$ is deductive, $\phi \vdash (A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m) \ldots) \rightarrow B$ for any B. Hence $\phi \vdash' (A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m) \ldots) \rightarrow B$ for any B, and so $\Gamma' \vdash' B$ for all B. By regularity, then, $\Gamma' \vdash' \phi$ and so $\Gamma \vdash' \Delta$. This shows $\vdash \subseteq \vdash'$. Since no assumption distinguishes $\vdash'$ from $\vdash$, the same argument shows $\vdash' \subseteq \vdash$.

Proof of (a). Since L is deductive and regular there must be a deductive and regular consequence relation which determines L. Let $\vdash$ be such a relation. We first check that $\vdash_L$ is a consequence relation. Since $\vdash$ is a consequence relation, $A \vdash A$. Since $\vdash$ is deductive, $\phi \vdash A \rightarrow A$ and hence $A \rightarrow A$ is in L. By clause (i) of the definition of $\vdash_L$, then, we know that if $\Gamma \cap \Delta$ contains the formula A then $\Gamma \vdash_L \Delta$. If $\Gamma \vdash_L \Delta$ it follows immediately from Definition 1.26 that $\Gamma \cup \Psi \vdash_L \Delta \cup \Theta$. It remains to check the third condition. If Γ' is a finite set we use $\wedge(\Gamma')$ and $\vee(\Gamma')$ the set of all formulas $(A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m)$ and $(A_1 \vee \ldots \vee (A_{m-1} \vee A_m) \ldots)$ respectively such that the A_i's run through Γ. Now suppose that $\Gamma \vdash_L \Delta, Q$ and $Q, \Psi \vdash_L \Theta$. There are five cases:

(i) Clause (iii) holds between Γ and Δ, Q. Then there is some finite set $\Gamma' \subset \Gamma$ and some A in (Γ') such that for all formulas B, $A \rightarrow B$ is in L. But since $\Gamma' \subseteq \Gamma \cup \Psi$ this is all that is needed to show $\Gamma \cup \Psi \vdash_L \Delta \cup \Theta$.

(ii) Clause (ii) holds between Q, Ψ and Θ. Then there is a finite set $\Theta' \subseteq \Theta$ and a formula B in $\vee(\Theta)$ such that B is in L. This is sufficient to show that $\Gamma,\Psi \vdash_L \Delta,\Theta$.

(iiia) Clause (ii) holds between Γ and Δ,Q and clause (i) holds between Q, Ψ and Θ. Then there is a $\Delta' \subseteq \Delta \cup \{Q\}$ such that for some B in Δ', B is in L, and there are $\Psi' \subseteq Q,\Psi$ and $\Theta' \subseteq \Theta$ such that for some C in $\wedge(\Psi')$ and some D in $\vee(\Theta')$, $C \to D$ is in L. Since $\vdash$ determines L we have $\phi \vdash B$ and $\phi \vdash C \to D$. Since $\vdash$ is deductive we have $\phi \vdash \Delta'$ and $\Psi' \vdash \Theta'$. If Q does not appear in Δ' or Ψ' we can add it (clause ii, definition 1.19), getting $\phi \vdash \Delta' -\{Q\},Q$ $Q,\Psi' -\{Q\} \vdash \Theta'$. Therefore $\Psi' - \{Q\} \vdash \Theta', \Delta' - \{Q\}$, and hence $\phi \vdash E \to F$ where E and F are members of $\wedge(\Psi' - \{Q\})$ and $\vee((\Delta' - \{Q\}) \cup \Theta')$ respectively. Since $L = L(\vdash)$, $E \to F \in L$. And since $\Psi' - \{Q\} \subseteq \Psi$ and $(\Delta' - \{Q\} \cup \Theta') \subseteq \Delta \cup \Theta$ this means that $\Gamma,\Psi =_L \Delta, \Theta$.

(iiib) Clause (ii) holds between Γ and Δ,Q and clause (iii) holds between Q,Ψ and Θ. Then there is a $\Delta' \subseteq \Delta,Q$ and a B in $\vee(\Delta')$ such that B is in L. Furthermore, there is a $\Psi' \subseteq Q,\Psi$ such that for some A in $\wedge(\Psi')$ $A \to C$ is in L for all formulas C. Since $\vdash$ determines L, this means that $\phi \vdash B$ and $\phi \vdash A \to C$ for all C. Hence $\phi \vdash \Delta'$ and $\Psi' \vdash C$, for all C. Since $\vdash$ is regular this means $\Psi' \vdash \phi$. Proceeding as in (iiia) we get $\phi, \vdash \Delta'',Q$ and $Q,\Psi'' \vdash \phi$, where $\Delta'' \subseteq \Delta$ and $\Psi'' \subseteq \Psi$. Hence $\Psi'' \vdash \Delta''$. Since $\vdash$ is deductive we get $\phi \vdash E \rightarrow F$ where E is now in $\wedge(\Psi'')$ and F in $\vee(\Delta'')$. This means that $E \rightarrow F$ is in L and therefore that $\Gamma,\Psi \vdash \Theta, \Delta$.

(iv) Clause (iii) holds between Q, Ψ and Θ and clause (i) holds between Γ and Δ,Q. Then there are finite sets $\Psi' \subseteq Q,\Psi$ $\Gamma' \subseteq \Gamma$, and $\Delta' \subseteq \Delta,Q$ such that for some formulas A in $\wedge(\Gamma')$, B in $\vee(\Delta')$ and C in $\wedge(\Psi')$, $A \rightarrow B$ is in L and for all formulas D, $C \rightarrow D$ is in L. As before, this implies $\Gamma' \vdash \Delta'$ and $\Psi' \vdash \phi$ which, in turn, implies $\Gamma' \vdash \Delta'',Q$ $Q,\Psi'' \vdash \phi$. Thus $\Gamma',\Psi'' \vdash \Delta''$ so $E \rightarrow F$ is in L where E and F are now members of $\wedge(\Gamma',\Psi'')$ and $\vee(\Delta'')$ respectively. This shows that $\Gamma,\Psi \vdash_L \Delta,\Theta$.

(v) Clause (i) holds between Γ and Δ,Q and also between Q, Ψ and Θ. Then there are finite sets $\Gamma' \subseteq \Gamma$; $\Delta' \subseteq \Delta,Q$; $\Psi' \subseteq Q,\Psi$; and $\Theta' \subseteq \Theta$ and formulas A, B, C, D in $\wedge(\Gamma')$, $\vee(\Delta')$, $\wedge(\psi')$, $\vee(\Theta')$ respectively such that $A \rightarrow B$ and $C \rightarrow D$ are both in L. This means that $\Gamma' \vdash \Delta'$ and $\Psi' \vdash \Theta'$. As before, this implies that $\Gamma' \vdash \Delta'',Q$ and $Q,\Psi'' \vdash \Theta'$, and therefore that $\Gamma',\Psi'' \vdash \Delta'', \Theta'$. Since $\vdash$ is deductive, $\phi \vdash E \rightarrow F$ where E and F are now members of $\wedge(\Gamma',\Psi'')$ and $\vee(\Delta'', \Theta')$ respectively. Thus $\Gamma,\Psi \vdash_L \Delta, \Theta$.

We now check that $\vdash_L$ is deductive. $\Gamma, A, B \vdash_L \Delta$ iff either (i), (ii) or (iii) of Definition 1.26 holds between Γ, A, B and Δ. If (ii) holds then (ii) still holds between $\Gamma, A \wedge B$ and Δ. If (i) or (ii) then there is a $\Gamma' \subseteq \Gamma, A, B$ and a $\Delta' \subseteq \Delta$ such that for some C in Γ' either $C \rightarrow D$ is in L for all D or $C \rightarrow E$ is in L for some E in Δ'. Hence either $\phi \vdash C \rightarrow D$ for all D or $\phi \vdash C \rightarrow E$. Since $\vdash$ is deductive and regular this means that $\Gamma' \vdash \Delta'$ or $\Gamma' \vdash \phi$. Expanding the left side if necessary, we get $\Gamma'', A, B \vdash \Delta'$ or $\Gamma'', A, B \vdash \phi$ where Γ'' does not contain A or B. Since $\vdash$ is deductive, $\Gamma'', A \wedge B \vdash \Delta'$ or $\Gamma'', A \wedge B \vdash \phi$ and therefore $\phi \vdash E \rightarrow F$ or $\phi \vdash E \rightarrow G$ for all G, where E is in $\wedge(\Gamma'', A \wedge B)$ and F is in $\vee(\Delta'')$. Thus $E \rightarrow F$ is in L or, for all G, $E \rightarrow G$ is in L. This shows $\Gamma, A \wedge B \vdash_L \Delta$. The other direction of this clause and the other clauses of Definition 1.20 are all handled similarly.

We now check that $\vdash_L$ is regular. Suppose first that $\Gamma \vdash_L B$ for every sentence B. Then for all B there is a set $\Gamma_B \subseteq \Gamma$ such that for some A_B in Γ_B, $A_B \rightarrow B$ is a member of L. Therefore $\phi \vdash A_B \rightarrow B$ for all B and, since $\vdash$ is deductive, $\Gamma_B \vdash B$ for all B. Therefore $\Gamma \vdash B$ for all B. But we are assuming that $\vdash$ is regular so this implies that $\Gamma \vdash \phi$. Furthermore, $\vdash$ is finitary so there is some finite set $\Gamma' \subseteq \Gamma$ such that $\Gamma' \vdash \phi$. Since $\vdash$ is deductive $A \vdash \phi$ where A is a member of $\wedge(\Gamma')$. But if $A \vdash \phi$ then by expansion $A \vdash B$ for all B. Hence $\phi \vdash A \rightarrow B$ for all B and, since $\vdash$ determines L, $A \rightarrow B$

is in L for all B. This is exactly what is needed to show $\Gamma \vdash_L \phi$. Next suppose that $B \vdash_L \Delta$ for every formula B. Then for all B there is a finite subset $\Delta_B \subseteq \Delta$ such that for some C_B in $v(\Delta_B)$, $B \to C_B$ is in L. Proceeding as before we get that $\phi \vdash \Delta$. But $\vdash$ is finitary, so there is a finite set $\Delta' \subseteq \Delta$ such that $\phi \vdash \Delta'$. Since L is deductive we have $\phi \vdash C$ where C is in $v(\Delta')$. Therefore C is in L. Again, this is all that is needed to show $\phi \vdash_L \Delta$.

The facts that $\vdash_L$ is finitary and $\vdash_L$ determines L follow immediately from Definition 1.23 so the proof of (a) is complete.

According to Lemma 1.7, if we knew $\vdash$ to be a regular, finitary, deductive consequence relation, we could replace talk about the consequence relation $\vdash$ by talk about the logic $L(\vdash)$ without fear of losing information. For it would always be possible to reconstruct $\vdash$ from $L(\vdash)$.

Definition 1.24. If L is a Boolean language and L is a deductive, regular logic on L, then $\Box$ is a Montague (Kripke, K4, S4, S5) connective with respect to L if it is a Montague (Kripke, K4, S4, S5) connective with respect to the finitary deductive regular consequence relation which determines L.

The notions described in Definitions 1.22 and 1.24 pertain to logics, yet the definitions were given in terms of consequence relations. It is also possible to characterize some of these notions directly.

<u>Theorem 1.1</u>. A set $\mathcal{L}$ of sentences of a language L is a deductive logic if and only if for all A, B, C, D, and Q in L, $\mathcal{L}$ contains:

1) $A \to (B \to A)$;

2) $A \to A$;

3) $(A \wedge B) \to (B \wedge A)$ and $(A \vee B) \to (B \vee A)$;

4) $(A \wedge (B \wedge C)) \to ((A \wedge B) \wedge C)$ and $(A \vee (B \vee C)) \to (A \vee B) \vee C)$;

5) $(A \wedge B) \to A$ and $A \to (B \vee A)$;

6) $A \to (A \wedge A)$ and $(B \vee B) \to B$;

7) $((A \wedge B) \to (C \vee D)) \to (A \to ((B \to C) \vee D)$;

8) $(A \to B) \to ((B \to C) \to (A \to C))$;

9) $(A \to (B \vee Q)) \to (((Q \wedge C) \to D) \to ((A \wedge C) \to (B \vee D)))$;

and $\mathcal{L}$ is closed under the rule:

MP) If A and $A \to B$ are in $\mathcal{L}$, so is B.

Furthermore, if $\mathcal{L}$ is deductive, then: $\mathcal{L}$ is regular if and only if A is in $\mathcal{L}$ whenever $B \to A$ is in $\mathcal{L}$ for every formula B.

<u>Theorem 1.2</u>. A set $\mathcal{L}$ of sentences of a language L is a classical logic if and only if, $\mathcal{L}$ is closed under MP and, for all formulas A, B, C, D in L, $\mathcal{L}$ contains 1) - 9) above and, in addition, it contains:

10) $(-A \to -B) \to (B \to A)$;

11) $(A \vee B) \to (-A \to B)$;

12) $(-A \to B) \to (A \vee B)$;

13) $(A \wedge B) \to -(A \to -B)$;

14) $-(A \to -B) \to (A \wedge B)$.

The proofs of Theorems 1.1 and 1.2 are omitted.

(The reader who knows some elementary logic may recognize 1-14 and 1-9 as axiom sets for the classical propositional calculus and the negation-free fragment of the classical propositional calculus, respectively.)

Definition 1.25. Let $\mathcal{L}$ be a logic on a Boolean language L. If $\Box$ is an n-ary connective ($n \geq 1$) of L then we define the following condition on $\mathcal{L}$:

$M_{\Box}$: For all $A_1,\ldots,A_n$ and all $B_1,\ldots,B_n$ in L, if $A_i \rightarrow B_i$ and $B_i \rightarrow A_i$ are in $\mathcal{L}$ for $1 \leq i \leq n$ then $\Box A_1\ldots A_n \rightarrow \Box B_1\ldots B_n$ and $\Box B_1\ldots B_n \rightarrow \Box A_1\ldots A_n$ are in $\mathcal{L}$.

If $\Box$ and $\boxminus$ are unary connectives then we will also consider the following conditions:

$N_{\Box}$: For all A in L, if A is in $\mathcal{L}$ so is $\Box A$;

$K_{\Box}$: For all A and B in L, $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$ is in $\mathcal{L}$;

$4_{\Box}$: For all A in L, $\Box A \rightarrow \Box\Box A$ is in $\mathcal{L}$;

$R_{\Box}$: For all A in L, $\Box A \rightarrow A$ is in $\mathcal{L}$;

$5_{\Box}$: For all A in L, $A \rightarrow \Box - \Box - A$ is in $\mathcal{L}$;

$T_{\Box,\boxminus}$: For all A in L, $A \rightarrow \Box - \boxminus - A$ and $A \rightarrow \boxminus - \Box - A$ are in $\mathcal{L}$.

Theorem 1.3.

i) If $\mathcal{L}$ is a deductive regular logic on a Boolean language L, then

a) $\Box$ is Montague with respect to L if and only if L satisfies $M_\Box$;

b) $\Box$ is normal-Kripke with respect to L if and only if L satisfies $N_\Box$ and $K_\Box$;

c) $\Box$ is normal-K4 with respect to L if and only if L satisfies $N_\Box$, $K_\Box$, and $4_\Box$;

d) $\Box$ is normal-S4 with respect to L if and only if L satisfies $N_\Box$, $K_\Box$, $4_\Box$, and $R_\Box$;

(ii) If L is classical, then

a) $\Box$ is S5 with respect to L if and only if L satisfies $N_\Box$, $K_\Box$, $4_\Box$, $R_\Box$, and $5_\Box$;

b) $\Box$, $\boxminus$ are a pair of tense connectives with respect to L if and only if L satisfies $N_\Box$, $N_\boxminus$, $K_\Box$, $K_\boxminus$, $4_\Box$, $4_\boxminus$, $R_\Box$, $R_\Box$ and $T_{\Box\boxminus}$.

<u>Proof</u>. Let $\vdash\ = \vdash_L$.

(i) (a) Suppose $\Box$ is Montague with respect to L_i, and for all i from 1 to n, $A_i \to B_i$ and $B_i \to A_i$ are in L. Then $\Box$ is Montague with respect to $\vdash$ and, since $\vdash$ is deductive, $A_i \dashv\vdash B_i$ for $1 \leq i \leq n$. Hence $\Box A_1 \ldots A_n \dashv\vdash \Box B_1 \ldots B_n$ and therefore $\phi \vdash \Box A_1 \ldots A_n \to \Box B_1 \ldots B_n$ and $\phi \vdash \Box B_1 \ldots B_n \to \Box A_1 \ldots A_n$. Since $\vdash$ determines L this shows that M is satisfied.

(a') Suppose L satisfies M and for all i from 1 to n, $A_i \dashv\vdash B_i$. Since $\vdash$ is deductive this means that $A_i \to B_i$ and $B_i \to A_i$ are in L for $1 \leq i \leq n$. By

$N_\Box$, $\Box B_1 \ldots B_n \to \Box A_1 \ldots A_n$ and $\Box A_1 \ldots A_n \to \Box B_1 \ldots B_n$ are in L. Hence $\Box A_1 \ldots A_n \dashv\vdash \Box B_1 \ldots B_n$.

(b) Suppose $\Box$ is normal-Kripke with respect to L. Then A in $L \Rightarrow \phi \vdash A \Rightarrow \Box(\phi) \vdash \Box A \Rightarrow \phi \vdash \Box A \Rightarrow \Box A$ is in L. So condition $N_\Box$ is satisfied. Furthermore, $A \to B$ $\vdash A \to B \Rightarrow A \to B, A \vdash B$ (since $\vdash$ is deductive) $\Rightarrow \Box(A \to B), \Box A \vdash \Box B$ (since $\Box$ is normal Kripke) $\Rightarrow$ $\Box(A \to B) \vdash \Box A \to \Box B \Rightarrow \phi \vdash \Box(A \to B) \to (\Box A \to \Box B) \Rightarrow$ $\Box(A \to B) \to (\Box A \to \Box B)$ is in L . So $K_\Box$ is also satisfied.

(b') Suppose L satisfies $N_\Box$ and $K_\Box$ and $\Gamma \vdash B$. If B is in L, then, by condition $N_\Box$, $\Box B$ is also in L. Hence $\phi \vdash \Box B$ and, by expanding if necessary, we get $\Box\Gamma \vdash \Box B$. If B is not in L then by the definition of $\vdash_L$ there must be a finite set $\{A_1,\ldots,A_n\} \subseteq \Gamma$ such that $(A_1 \wedge \ldots \wedge (A_{n-1} \wedge A_n)\ldots) \to B$ is in L. But in this case $\{A_1,\ldots,A_n\} \vdash B$ so since $\vdash$ is deductive, $\phi \vdash (A_1 \to \ldots \to (A_{n-1} \to (A_n \to B))\ldots)$. Therefore $A_1 \to \ldots \to (A_n \to B)\ldots$ is in L. By $N_\Box$, $\Box(A_1 \to \ldots \to (A_n \to B)\ldots)$ is also in L. But by K and the fact that L is deductive

$$(A_1 \to \ldots \to (A_n \to B)\ldots) \to (\Box A_1 \to (\Box A_2 \to \ldots \to (\Box A_n \to \Box B)\ldots))$$

is in L. Hence $\Box A_1 \to (\ldots \to (\Box A_n \to \Box B)\ldots)$ is in L, i.e., $\phi \vdash \Box A_1,\ldots,\Box A_n, \Box B$. So $\Box A_1,\ldots,\Box A_n \vdash \Box B$. Expanding if necessary, $\Box\Gamma \vdash \Box B$.

(c) Suppose $\Box$ is normal K4 with respect to $\vdash$. By (b), it suffices to show that L satisfies $4_\Box$. Since $\Box A \vdash \Box A$ and $\Box$ is normal K4 with respect to $\vdash$, $\Box A \vdash \Box\Box A$. Hence $\phi \vdash \Box A \rightarrow \Box\Box A$, so $\Box A \rightarrow \Box\Box A \in L$.

(c') Suppose L satisfies $N_\Box, K_\Box$, and $4_\Box$. By (b'), $\Box$ is normal-Kripke with respect to $\vdash$. Now suppose $\Box\Gamma \vdash A$. If $A \in L$, then by $N_\Box$, $A \in L$. Hence $\phi \vdash \Box A$ so $\Box\Gamma \subseteq \Box A$. If $A \notin L$, then there must be a set $\{\Box B_1, \ldots, \Box B_n\} \subseteq \Gamma$ such that $(\Box B_1 \wedge \ldots \wedge (\Box B_{n-1} \wedge \Box B_n) \ldots) \rightarrow A$ is in L. Hence $\Box B_1, \ldots, \Box B_n \vdash A$. Since $\Box$ is normal-Kripke with respect to $\vdash$, $\Box B_1, \ldots, \Box B_n \vdash \Box A$. But $\Box B_i \rightarrow \Box\Box B_i$ in in L for all B_i so $\Box B_i \vdash \Box\Box B_i$ $(1 \leq i \leq n)$. Therefore $\Box B_1, \Box B_2, \ldots, \Box B_n \vdash \Box A$. Hence $\Box\Gamma \vdash \Box A$.

(d) Suppose $\Box$ is normal-S4 with respect to $\vdash$. By (c) L satisfies $N_\Box$, $K_\Box$, $4_\Box$. Since $A \vdash A$ and $\Box$ is normal S4 with respect to $\vdash$, $\Box A \vdash A$. Hence $\Box A \rightarrow A$ is in L.

(d') Suppose L satisfies $N_\Box$, $K_\Box$, $4_\Box$, and $R_\Box$. By (c'), $\Box$ is normal-K4 with respect to $\vdash$. Now suppose $A, \Gamma \vdash \Delta$. Since $\Box A \rightarrow A$ is in L we have $\Box A \vdash A$. Therefore $\Box A, \Gamma \vdash \Delta$.

(ii) (a) Suppose $\Box$ is normal-S5 with respect to $\vdash$. By (d), L satisfies $N_\Box$, $K_\Box$, $4_\Box$, and $R_\Box$. Since

$\Box -A \vdash \Box -A$ and L is assumed to have classical negation $\phi \vdash \Box -A, - \Box -A$. Since $\Box$ is normal-S5 with respect to $\vdash$, $\phi \vdash \Box -A, \Box - \Box -A$. Hence $- \Box -A \vdash \Box - \Box -A$. But we also know that $\Box -A \rightarrow -A$ is in L so $\Box -A \vdash -A$. This means that $A \vdash - \Box -A$. Thus $A \vdash \Box - \Box -A$ and $A \rightarrow \Box - \Box -A$ is in L.

(a') Suppose L has classical negation and L satisfies N , K , 4 , R and S . By (d') $\Box$ is normal S4 with respect to $\vdash$. Suppose $\Box\Gamma \vdash \Box\Delta, A$. Since $\vdash$ is finitary there are formulas $A_1,\ldots,A_m$ in Γ and $B_1,\ldots,B_n$ in Δ such that $\Box A_1,\ldots,\Box A_m \vdash \Box B_1,\ldots,\Box B_n, A$. But, since $\vdash$ is normal-K4 $\Box B \vdash \Box\Box B$ for all B. Therefore $\Box A_1,\ldots,\Box A_m \vdash \Box\Box B_1,\ldots,\Box\Box B_n, A$. Since L has classical negation, $\Box A_1,\ldots,\Box A_m, -\Box\Box B_1,\ldots,-\Box\Box B_n \vdash A$, and $\Box B \vdash -- \Box B$ for all B. Furthermore $\Box$ is Kripke, so $\Box\Box B \vdash \Box --\Box B$ and thus $-\Box --\Box B - -\Box\Box B$. By n applications of cut, $\Box A_1,\ldots,\Box A_m, -\Box -- \Box B_1,\ldots,-\Box --\Box B_n \vdash A$. Since $\Box$ is Kripke, $\Box\Box A_1,\ldots,\Box\Box A_m, \Box-\Box--\Box B_1,\ldots, \Box - \Box --\Box B_n \vdash A$. Furthermore $\Box C \vdash \Box\Box C$ for all C. Hence $\Box A_1,\ldots,\Box A_m, \Box-\Box-- \Box B_1,\ldots, \Box - \Box -- \Box B_n \vdash \Box A$. Now L satisfies $5_{\Box}$, so $-\Box B \vdash \Box -\Box -- \Box B$ for all B, and, by n more cuts, $\Box A_1,\ldots, \Box A_m, - \Box B_1,\ldots, - \Box B_m \vdash \Box A$. In addition L has classical negation, so $\Box A_1,\ldots,\Box A_m \vdash \Box B_1,\ldots,\Box B_m,\Box A$, from which we obtain $\Box\Gamma \vdash \Box\Delta, \Box A$.

(b) Suppose $(\Box, \boxminus)$ is a pair of tense connectives with respect to $\vdash$. By (d) L satisfies $K_{\Box}$, $K_{\boxminus}$, $4_{\Box}$, $4_{\Box}$, $R_{\Box}$, and $R_{\boxminus}$. Since $\vdash$ has classical negation we know that $\phi \vdash \Box -A, \; -\Box -A$ and $\phi \vdash \Box -A, \; -\Box -A$. Hence $\phi - \Box -A, \; \boxminus - \Box -A$ and $\phi \vdash \boxminus -A, \; \Box - \boxminus -A$, from which we obtain $-\Box -A \vdash \boxminus - \Box -A$ and $-\boxminus -A - \Box - \boxminus -A$. Furthermore, since $\Box$ and $\boxminus$ are both normal S4 with respect to $\vdash$, $\Box -A \vdash -A$ and $\boxminus -A \vdash -A$. So $A - \boxminus - \Box -A$ and $A - \Box - \boxminus -A$. Thus L contains both $A \rightarrow \boxminus - \Box -A$ and $A \rightarrow \Box - \boxminus -A$.

(b') Suppose L satisfies $K_{\Box}$, $K_{\boxminus}$, $4_{\Box}$, $4_{\boxminus}$, $R_{\Box}$, $R_{\boxminus}$ and $T_{\Box\boxminus}$. By (d'), both $\Box$ and $\boxminus$ are normal-S4 with respect to $\vdash$. Now suppose $\boxminus\Gamma \vdash \Box\Delta, A$. By using the same reasoning as in (e') we can get $\boxminus A_1, \ldots, \boxminus A_m, \; -\Box -- \Box B_1, \ldots, \; -\Box -- \Box B_n \vdash A$ for some $A_1, \ldots, A_m$ in Γ and some $B_1, \ldots, B_n$ in Δ. Since $\boxminus$ is Kripke, this means, $\boxminus\boxminus A_1, \ldots, \boxminus\boxminus A_m$, $\boxminus - \Box -- \Box B_1, \ldots, \boxminus - \Box -- \Box B_n - \boxminus A$. Furthermore $\boxminus A - \boxminus\boxminus A$ by $4_{\Box}$ and $-\Box B_1 \vdash \boxminus - \Box -- \Box B_1$ by $T_{\Box\boxminus}$. So $\boxminus A_1, \ldots, \boxminus A_m, \; -\Box B_1, \ldots, \; -\Box B_n \vdash \boxminus A$. Hence $\boxminus A_1, \ldots, \boxminus A_m \vdash \boxminus A, \Box B_1, \ldots, \Box B_n$. An analagous argument shows that $\Box A_1, \ldots, \Box A_m \vdash \boxminus B_1, \ldots, \boxminus B_n \; \Box A$. Thus $(\Box, \boxminus)$ is a pair of tense connectives with respect to $\vdash$.

G. Completeness and Axiomatization

Throughout this section we assume that L is a language, $\mathcal{M}$ is a class of models suitable for L, $\vdash$ is a consequence relation on L, and $\mathcal{L}$ is a logic on L.

Definition 1.26. $\vdash$ is sound with respect to $\mathcal{M}$ if $\vdash \subseteq \models_{\mathcal{M}}$. $\mathcal{L}$ is sound with respect to $\mathcal{M}$ if $\phi \models_{\mathcal{M}} A$ whenever A is in $\mathcal{L}$.

Definition 1.27. $\mathcal{M}$ is weakly sufficient for $\vdash$ if for all formulas A,B $A \vdash B$ whenever $A \models_{\mathcal{M}} B$. $\mathcal{M}$ is sufficient for $\vdash$ if for all finite sets Γ,Δ, $\Gamma \vdash \Delta$ whenever $\Gamma \models_{\mathcal{M}} \Delta$. $\mathcal{M}$ is strongly sufficient for $\vdash$ if for all sets Γ,Δ, $\Gamma \vdash \Delta$ whenever $\Gamma \models_{\mathcal{M}} \Delta$.

Lemma 1.8. If $\vdash$ is finitary then it is strongly sufficient for $\mathcal{M}$ if and only if it is sufficient for $\mathcal{M}_1$. If $\vdash$ is deductive then it is sufficient for $\mathcal{M}$ if and only if it is weakly sufficient for $\mathcal{M}$.

Definition 1.28. $\vdash$ is weakly complete (complete, strongly complete) for $\mathcal{M}$ if for all formulas (finite sets of formulas, sets of formulas) X and Y , $X \models_{\mathcal{M}} Y$ iff $X \vdash Y$. $\mathcal{L}$ is complete for $\mathcal{M}$ if $A \in \mathcal{L}$ if and only if $\phi \models_{\mathcal{M}} A$.

Notice that to prove completeness (weak completeness, strong completeness) it is sufficient to prove soundness and sufficiency (weak sufficiency, strong sufficiency).

Theorem 1.4 (Soundness). Let $\mathcal{M}$ be a normal class of models suitable for L. Then

a) Every connective of L is Montague with respect to $\models_{\mathcal{M}}$.

b) If every member of $\mathcal{M}$ is partially classical then $\models_{\mathcal{M}}$ is partially classical.

c) If, for all M in $\mathcal{M}$, $\bar{\Box}_M$ is a point-to-point (transitive point-to-point; transitive and reflexive point-to-point; transitive, reflexive and symmetric point-to-point) relation then $\Box$ is a normal-Kripke (normal-K4, normal-S4, normal-S5) connective with respect to $\models_{\mathcal{M}}$.

d) If, for all M in $\mathcal{M}$, $\bar{\Box}_M$ and $\bar{\boxminus}_M$ are transitive, reflexive, point-to-point relations such that $\bar{\Box}_M$ is the converse of $\bar{\boxminus}_M$, then $(\Box, \boxminus)$ is a pair of tense connectives, with respect to $\models_{\mathcal{M}}$.

e) If L is a predictive language and $\mathcal{M}$ is assignment-unrestricted, then $\models_{\mathcal{M}}$ has classical quantifiers.

f) If $\mathcal{M}$ is valuation-unrestricted, then $\models_{\mathcal{M}}$ is substitution closed.

Proof. Let $\models \; = \; \models_{\mathcal{M}}$.

a) Suppose $\Box$ is n-ary, $A_i \dashv\models B_i$ for $1 \leq i \leq n$ and $M = (W, 0, \ldots, \bar{\Box}, \ldots, V)$ is a model in $\mathcal{M}$.

Case (i): $\bar{\Box}$ is an n-ary neighborhood relation. Then $M \models \Box A_1 \ldots A_n$ iff $\bar{\Box}\, o, |A_1|_M, \ldots, |A_n|_M$. But $(M,u) \models A_i$ iff $M^u \models A_i$ and, since $\mathcal{M}$ is normal, each M^u is also in $\mathcal{M}$. Since $A_i \dashv\models B_i$ we know that $|A_i|_M = \{u : M^u \models A_i\} = \{u : M^u \models B_i\} = |B_i|_M$. Hence $\bar{\Box}\, o, |A_1|_M, \ldots, |A_n|_M$ iff

$\bar{\Box}\, o, |B_1|_M, \ldots, |B_n|_M$ iff $M \models \Box B_1 \ldots B_n$.

Case (ii): $\Box$ is unary and $\bar{\Box}$ is a point-to-point relation. Then $M \models \Box A$ iff, for all w in W, $o\,\bar{\Box}\,w$ implies $M^w \models A$. But since M^w is in $\mathcal{M}$ and $A \dashv\models B$ this is equivalent to the conditions for all w in W, $o\,\bar{\Box}\,w$ implies $M^w \models B$. Thus in either case $M \models \Box A_1 \ldots A_n$ iff $M \models \Box B_1 \ldots B_n$.

b) Suppose that every M in $\mathcal{M}$ is partially classical. Using Lemma 1.1 it is routine to verify CC, CD, CI, CN of Definition 1.17. We do half of clause CN as an example. $\Gamma \vdash A,\Delta$ iff for all M in $\mathcal{M}$, if $M \models \Gamma$ then there is a B in $\{A\} \cup \Delta$ such that $M \models B$, i.e., iff, if $M \models \Gamma$ and $M \not\models A$, then there is a B in Δ such that $M \models B$. But since M is partially classical, this just means $\Gamma, -A \vdash B$.

c) i. Suppose that for all M in $\mathcal{M}$ $\bar{\Box}_M$ is a point-to-point relation, that $\Gamma \models A$ and that $M \models \langle W, 0, \ldots, \bar{\Box}, \ldots, V\rangle$ is a model in $\mathcal{M}$ such that $M \models \Box\Gamma$. Then, for all B in Γ and all w in W, of $o\,\bar{\Box}\,w$ then $(M,w) \models B$, i.e., $o\,\bar{\Box}\,w$ implies $(M,w) \models \Gamma$. But since $(M,w) \models \Gamma$ and M^w is in $\mathcal{M}$, this means $o\,\bar{\Box}\,w$ implies $(M,w) \models A$, i.e., that $(M,o) \models \Box A$.

ii). Suppose that for all M in $\mathcal{M}$, $\bar{\Box}_M$ is transitive and that $\Box\Gamma \models A$. We know by i) above that $\Box\Box\Gamma \models \Box A$. Let $M = \langle W, 0, \ldots, \bar{\Box}, \ldots, V\rangle$ be a model in $\mathcal{M}$ such that $M \models \Box\Gamma$. Then for all u in W, if $o\,\bar{\Box}\,u$ then $M^u \models \Gamma$. But $\bar{\Box}$ is transitive, so for all w in W the following holds: if

$o \bar{\Box} u$, then, for all v in W, if $u \bar{\Box} v$ then $o \bar{\Box} v$ (and therefore $M^v \models \Gamma$). This is just what is needed to establish that $M \models \Box\Box\Gamma$. Hence $M \models \Box A$.

iii). Suppose for all M in $\mathcal{M}$, $\bar{\Box}_M$ is transitive and reflexive and $A, \Gamma \models \Delta$. Let $M = (W, 0, \ldots, \bar{\Box}, \ldots, V)$ be a model in $\mathcal{M}$ such that $M \models A, \Gamma$. Then for all u in W, $o \bar{\Box} u$ implies $M^u \models A$. But since $\bar{\Box}$ is reflexive, $o \bar{\Box} o$, and hence $M \models A$. Since $A, \Gamma \models \Delta$ there must be a B in Δ such that $M \models B$.

iv). Suppose that for all M in $\mathcal{M}$, $\bar{\Box}_M$ is an equivalence relation and that $\Box\Gamma \models \Box\Delta, A$. Let $M = (W, 0, \ldots, \bar{\Box}, \ldots, V)$ be a model in $\mathcal{M}$ such that $M \models \Box\Gamma$. Then for some B in $\Box\Delta \cup \{A\}$, $M \models B$. We must show that if $M \not\models \Box A$ then for some B in $\Box\Delta$, $M \not\models B$. Now if $M \models \Box A$, then $M^u \not\models A$ for some u such that $o \bar{\Box} u$. Since $M \models \Box\Gamma$ and $\bar{\Box}$ is transitive we know $M^u \models \Box C$ for some $\Box C$ in $\Box\Delta$. From the transitivity of $\bar{\Box}$, it follows that $M^u \models \Box\Box C$ and, since $\bar{\Box}$ is symmetric, we know $u \bar{\Box} o$. Hence $M^o \models \Box C$.

d) Suppose, for all M in $\mathcal{M}$, $\bar{\Box}_M$ and $\bar{\boxminus}_M$ are transitive, reflexive and $\bar{\Box}_M$ is the converse of $\bar{\boxminus}_M$. By c) ii. above we know that $\Box$ and $\boxminus$ are K4 connectives with respect to $\models_{\mathcal{M}}$. Now suppose $\Box\Gamma \models \boxminus\Delta, A$ and let $M = (W, 0, \ldots, \bar{\Box}, \bar{\boxminus}, \ldots V)$ be a model in $\mathcal{M}$ such that $M \models \Box\Gamma$. Then there is a B in $\boxminus\Delta \cup \{A\}$ such that $M \models B$. We must show that, if there is no such B in $\boxminus\Delta$, then $M \models \Box A$. If this weren't true then, for some u such that $o \bar{\Box} u$ $M^u \not\models A$. Since $M \models \Box\Gamma$ and

$\overline{\Box}$ is transitive, we know $M^u = \Box\Gamma$. But $\Box\Gamma \models \boxminus\Delta, A$ so $M^u = \boxminus C$ for some $\boxminus C$ in $\boxminus\Delta$. From the transitivity of $\overline{\boxminus}$, it follows that $M^u \models \boxminus\boxminus C$ and, since $\overline{\boxminus}$ is the converse of $\overline{\Box}$, $u \,\overline{\overline{\boxminus}}\, o$. Hence $M^o \models \boxminus C$. But we were assuming that there was no such C. The case for $\boxminus$ is handled similarly.

e) i) Suppose $\Gamma \models \Delta, A(x)$, y does not occur free in $A(x)$ and $M = \langle W, 0, D, P, \overline{C}, a \rangle$ is a model in $\mathcal{M}$ such that $M \models \Gamma$ and for all B in Δ $M \not\models B$. We must show $M = \exists y A(y)$. Since $\Gamma \models \Delta, A(x)$, we know $M \models A(x)$. Let $M' = \langle W, 0, D, P, \overline{C}, a' \rangle$ where $a'(z) = a(z)$ for $z \neq y$ and $a'(y) = a(x)$. Then by Lemma 1.1 (5), $M' \models A(y)$. Hence $M \models \exists y A(y)$.

ii) Suppose $\Gamma, A(x) \models \Delta$ and $M = \langle W, 0, D, P, \overline{C}, a \rangle$ is a model in $\mathcal{M}$ such that $M \models \Gamma \cup \{\exists y A(y)\}$ and that x does not occur free in Γ, $\exists y(A_y)$ or Δ. $M \models \exists y A(y)$ implies there is an M' such that $M' \models A(y)$ and $M' = \langle W, 0, D, P, \overline{C}, a' \rangle$ where $a'(z) = a(z)$ for all $z \neq y$. Let $M'' = \langle W, 0, D, P, \overline{C}, a'' \rangle$ where $a''(z) = a(z)$ for $z \neq x$ and $a''(x) = a'(y)$. Since x is not free in Γ $M'' \models \Gamma$ (see Lemma 1.1g). Furthermore $M'' \models A(x)$ since $M' = A(y)$. But M'' is in $\mathcal{M}$, so there must be a C in Δ such that $M'' \models C$. But, since x is not free in Δ this means $M \models C$.

iii) The cases for $\forall x$ are treated similarly.

f) Suppose $\Gamma \vdash \Delta$ and Γ' and Δ' are the results of substituting the sentence A for all occurrences of the sentence letter p

in Γ and Δ. If M is a member of $\mathcal{M}$ such that $M \models \Gamma'$, let M' be the same as M except that $V_{M'}(P) = |A|_M$. Then $M' \models \Gamma$. Hence there is some B in Δ such that $M' \models B$. This implies that $M \models B'$ where B' is the result of substituting A for all occurrences of p in B.

Before turning to the matter of completeness we define some notions that will be useful when we come to construct models.

Definition 1.29.[11] A theory in L is a pair $\langle \Gamma, \Delta \rangle$ such that $\Gamma \subseteq L$ and $\Delta \subseteq L$. $\langle \Gamma, \Delta \rangle$ is finite if $\Gamma \cup \Delta$ is finite.

Definition 1.30. $\langle \Gamma, \Delta \rangle$ is $\vdash$consistent if $\Gamma \nvdash \Delta$.

Definition 1.31. $\langle \Gamma', \Delta' \rangle$ is an extension of $\langle \Gamma, \Delta \rangle$ if $\Gamma \subseteq \Gamma'$ and $\Delta \subseteq \Delta'$. A theory T is a proper extension of a theory T' if T is an extension of T', but T' is not an extension of T.

Definition 1.32. A theory is maximal $\vdash$consistent if it has no proper $\vdash$consistent extensions.

Lemma 1.9 (Lindenbaum's Lemma). If $\vdash$ is finitary, then every $\vdash$ consistent theory has a maximal $\vdash$ consistent extension.

Proof. Suppose $\langle \Gamma, \Delta \rangle$ is an $\vdash$ consistent theory in L. Let $A_1, A_2, \ldots$ be an enumeration of the formulas in L. Let $\Gamma_0 = \Gamma$,

---

[11] The terminology here follows [Gabbay, 1974].

$\Delta_0 = \Delta$. If $\langle \Gamma_i \cup \{A_{i+1}\}, \Delta_i \rangle$ is consistent, let $\Gamma_{i+1} = \Gamma_i \cup \{A\}$, and let $\Delta_{i+1} = \Delta_i$. Otherwise let $\Gamma_{i+1} = \Gamma$ and $\Delta_{i+1} = \Delta_i \cup \{A\}$. Let $\Gamma' = \bigcup_{i<w} \Gamma_i$ and let $\Delta' = \bigcup_{i<w} \Delta_i$. Claim: $\langle \Gamma', \Delta' \rangle$ is a maximal $\vdash$-consistent theory. If it is not consistent, then there are finite sets $\Gamma'' \subseteq \Gamma'$ and $\Delta'' \subseteq \Delta'$ such that $\Gamma'' \vdash \Delta''$. Hence there is some i such that $\Gamma_i \vdash \Delta_i$. Let m be the least such i (since $\langle \Gamma_0, \Delta_0 \rangle$ is $\vdash$-consistent, $m > 0$). From our construction we know that A_m is in either Γ_m or Δ_m. Furthermore it would only be in Γ_m if $\Gamma_m \nvdash \Delta_m$ so it must be the case that A is in Δ_m and $\Gamma_{m-1}, A_m \vdash \Delta_{m-1}$. But since $\Delta_m = \Delta_{m-1} \cup \{A\}$ and $\Gamma_m = \Gamma_{m-1}$, $\Gamma_{m-1} \vdash \Delta_{m-1}, A$. By an application of cut, $\Gamma_{m-1} \vdash \Delta_{m-1}$ contradicting our assumption that m was the first inconsistent stage. This proves consistency. Now suppose $\Gamma' \subseteq \Psi$ and $\Delta' \subseteq \Theta$ and $\Psi \vdash \Theta$. Let A be any formula in Ψ. The A must be one of the A_i's on the list, say A_m, and therefore A must be in Γ_{m-1} or Δ_{m-1}. But $\Delta_{m-1} \subseteq \Delta' \subseteq \Theta$ so A could not be in Δ' without violating the consistency of $\langle \Psi, \Theta \rangle$. Hence $\Psi \subseteq \Gamma'$. Similarly, if B is in Θ it must be in Δ_m for some m and hence in Δ'. Therefore (Γ', Δ') has no proper extensions.

<u>Lemma 1.10</u>. Suppose $\langle \Gamma, \Delta \rangle$ is maximal $\vdash$ consistent.

(a) For all A in L, A is either in Γ or Δ, but not both.

(b) If $\vdash$ is partially classical then:

If L contains $\wedge$, $A \wedge B$ is in Γ if and only if both A and B are in Γ; if L contains $\vee$, $A \vee B$ is in Γ if and only if either A or B is in Γ; if L contains $\rightarrow$ then

$A \to B$ is in Γ if and only if either A is not in Γ or B is in Γ. If L contains $-$ then $-A$ is in Γ if and only if A is not in Γ.

Proof. (a) Suppose A was in neither Γ nor Δ. Then we must have $\Gamma, A \vdash \Delta$ and $\Gamma \vdash \Delta, A$, (otherwise $\langle \Gamma \cup \{A\}, \Delta \rangle$ or $\langle \Gamma, \Delta \cup \{A\} \rangle$ would be a proper $\vdash$ consistent extension of $\langle \Gamma, \Delta \rangle$). By cut, $\Gamma \vdash \Delta$, which contradicts the consistency of (Γ, Δ).

(b) We prove two cases and leave the others to the reader.

(i) $A \wedge B \vdash A \wedge B$ and $\vdash$ deductive implies $A, B \vdash A \wedge B$. Hence if A and B are in Γ, $A \wedge B$ can't be in Δ. By (a), then A, B in Γ implies $A \wedge B$ in Γ, $A, B \vdash A$ and $A, B \vdash B$ imply $A \wedge B \vdash A$ and $A \wedge B \vdash B$. Hence $A \wedge B$ in Γ implies niether A nor B can be in Δ, i.e., both A and B must be in Γ.

(ii) $A \to B \vdash A \to B$ implies $A \to B, A \vdash B$. Hence if $A \to B$ and A are in Γ, B must also be in Γ, i.e., if $A \to B$ is in Γ, then either A is not in Γ or B is. $A, B \vdash B$ and $A \vdash A, B$ imply $B \vdash A \to B$ and $\phi \vdash A, A \to B$, respectively. Hence if B is in Γ, $A \to B$ must be in Γ and if A is in n ; $\vdash$ in Γ then $A \to B$ must be in Γ.

By part (a) of the above we can identify a maximal consistent theory (Γ, Δ) with its first coordinate Γ; for Δ must always equal $L - \Gamma$. We shall do this frequently in what follows. If (Γ, Δ) is such a theory and $A \in \Gamma$, we call Γ a maximal $\vdash$ consistent set containing A.

So far we have not distinguished between propositional and predicate languages. For the next few pages, however, we interrupt this uniform treatment, and assume until further notice that L is a propositional language.

<u>Definition 1.33a</u>. If A is in L, then $\underline{\|A\|_{\vdash}}$ is the set of all maximal -consistent sets containing A.

<u>Lemma 1.11a</u>. $\|A\|_{\vdash} \subseteq \|B\|_{\vdash}$ if and only if $A \vdash B$.

<u>Proof</u>. If $A \vdash B$, T is maximal consistent and A is in T then clearly B must be in T. If $A \nvdash B$, then $\langle \{A\},\{B\} \rangle$ is $\vdash$ consistent and has a maximal consistent extension T. But T will contain A and not B.

<u>Definition 1.34a</u>. If $\Box$ is a Kripke connective with respect to $\vdash$, then $\underline{\|\Box\|_{\vdash}}$ is the binary relation which holds between maximal $\vdash$ consistent sets u,v if and only if, for all sentences A, A is in v whenever $\Box A$ is in u. If $\Box$ is an n-ary Montague connective which is not a Kripke connective, then $\underline{\|\Box\|_{\vdash}}$ is the relation which holds between a maximal $\vdash$ consistent set w and an n-tuple $\langle u_1,\ldots,u_n \rangle$ of sets of maximal $\vdash$ consistent sets if and only if, for some formulas $A,\ldots, A_n$ in L, $u_1 = \|A_1\|_{\vdash},\ldots,$ $u_n = \|A_n\|_{\vdash}$ and $A_1\ldots A_n$ is in w.

<u>Definition 1.35a</u>. A <u>canonical</u> <u>modal</u> <u>model</u> <u>for</u> $\underline{\vdash}$ is a modal model $\langle W,0,\bar{C},V \rangle$ suitable for L such that w is the set of all maximal $\vdash$ consistent theories, $\bar{C} = \{\|\Box\|_{\vdash} : \Box \text{ is a connective of } L\}$ and $V(p) = \{w \in W : p \in w\}$.

Lemma 1.12a. If M is a canonical modal model for $\vdash$, then $(M,w) \models A$ if and only if $A \in w$.

Proof. Let $M = \langle W, o, \bar{C}, V \rangle$. We prove the lemma by induction on the length of A.

• If A is atomic, $(M,w) \models A$ iff $w \in V(A)$ iff $A \in w$.

• If $A = \Box A_1 \ldots A_n$ and $\Box$ is not Kripke then $(M,w) \models A$ iff $\bar{\Box}\, w, |A_1|_M, \ldots, |A_n|_M$. But each A_j is shorter than A, so by the induction hypothesis $(M,v) \models A_j$ iff $A_j \in v$. Hence $|A_j|_M = \|A_j\|_\vdash$ and $(M,w) \models A$ iff $\bar{\Box}\, w, \|A_1\|_\vdash, \ldots, \|A_n\|_\vdash$. Now, if $(M,w) \not\models A$ then not $\bar{\Box}\, w, \|A_1\|_\vdash, \ldots, \|A_n\|_\vdash$. Since $\bar{\Box} = \|\Box\|_\vdash$, this means $\Box A_1 \ldots A_n$ can't be in w. On the other hand, if $(M,w) \models A$ then $\bar{\Box}\, w, \|A_1\|_\vdash, \ldots, \|A_n\|_\vdash$, so there are sentences $B_1, \ldots, B_n$ such that for $1 \leq i \leq n$ $\|B_i\|_\vdash = \|A_1\|_\vdash$ and $\Box B_1 \ldots B_n$ is in w. So by Lemma 1.11a, $B_i \dashv\vdash A_i$ for $1 \leq i \leq n$, and, since $\vdash$ is Montague, $\Box B_1 \ldots B_n \dashv\vdash \Box A_1 \ldots A_n$. Hence $\Box A_1 \ldots A_n$ must also be in w.

• If $\Box$ is Kripke and $(M,w) \not\models \Box B$, then there is some u such that $w \bar{\Box} u$ but $(M,u) \not\models B$. By induction hypothesis this means that $B \notin u$. So since $\bar{\Box} = \|\Box\|_\vdash$, $\Box B$ can't be in w. Conversely, suppose $B \notin w$ and let u be the theory $\langle \{C : C \in w\}, (B) \rangle$. u must be consistent, because otherwise we would have $\Gamma \vdash B$ for some $\Gamma . \subseteq \{C : C \in w\}$ and hence $\Box\Gamma \vdash \Box B$ (since $\Box$ is Kripke), thus violating the consistency of w.

• We can therefore extend u to a maximal $\vdash$ consistent theory u'. Clearly $w \|\Box\|_\vdash u'$ and $B \notin u'$. By induction hypothesis $(M,u') \not\models B$. Hence $(M,w) \not\models \Box B$.

We now want to prove the analogue of Lemma 1.12a for predicate languages. So we assume below that L is a predicate language.

Definition 1.36. A theory (Γ,Δ) in L is $\vdash$ saturated if it is maximal $\vdash$ consistent and for every formula A of the form $\exists xB(x)$ A is in Γ if and only if, for some variable y, $B(y)$ is in Γ.

Lemma 1.13. Let L' be the language obtained from L by adding a countable collection of new variables, and let $\vdash'$ be a finitary consequence relation on L' such that $\vdash' \cap (2^L \times 2^L) = \vdash$. Then every $\vdash$ consistent theory has a $\vdash'$ saturated extension.

Proof. Let (Γ,Δ) be a $\vdash$ consistent theory and let (Γ',Δ') be a maximal $\vdash$ consistent extension of (Γ,Δ). Let $A_1,A_2,\ldots$ be an enumeration of all the formulas of L' and let $u_1,u_2,\ldots$ be an enumeration of the variables in $L' - L$. If $x = u_i$ we say x has index i and if x is a variable of L we say it has index 0. We define the theory T^+ in stages: $T_0^+ = (\Gamma',\Delta')$. If A_m is not of the form $\exists xB(x)$ then T_{m+1}^+ is obtained by adding A_m to the left coordinate of T_m^+ if the result is $\vdash'$ consistent, and by adding it to the right side otherwise. If $A_m = \exists xB(x)$ then let $j = m$ plus the greatest number which is the index of a variable that occurs free in $A_1,\ldots,A_m$. T_{m+1}^+ is then obtained by adding $\exists xB(x)$ and $B(u_{j+1})$ to the left coordinate of T_m^+ if the result is $\vdash'$ consistent and by adding $\exists xB(x)$ to the right coordinate otherwise. Let T^+ be the union of all the T_i^+'s. It is easy to to check that T^+ is a $\vdash'$ saturated extension of (Γ,Δ).

Let L' be as in Lemma 1.13 and let $\vdash'$ be the relation which holds between two subsets Γ' and Δ' of L' if and only if there are $\Gamma \subseteq L$ and $\Delta \subseteq L$ such that $\Gamma \vdash \Delta$ and Γ', Δ' are obtained from Γ,Δ by uniformly substituting variables not in L for some or all of the variables in Γ,Δ. It is easy to check that i) $\vdash'$ is a consequence relation; ii) $\vdash' \cap (2^L \times 2^L) = \vdash$ and iii) $\vdash$ is finitary (deductive, classical regular) whenever $\vdash$ is. Also $\Box$ is Montague (normal-Kripke, normal-K4, normal-S4, normal-S5) with respect to $\vdash'$ whenever it is Montague (normal-Kripke, normal-K4, normal-S4, normal-S5) with respect to $\vdash$.

Definition 1.33b. Let $W_\vdash$ be the set of all $\vdash'$ saturated sets of sentences of L'. If A is in L, then $\|A\|_\vdash$ is the set of all w in $W_\vdash$ which contain A.

(Notice that the notation here is the same as that introduced in 1.33a. The intended interpretation will always be clear from context.)

Lemma 1.11b. $\|A\|_\vdash \subseteq \|B\|_\vdash$ if and only if $A \vdash B$.

Proof. If $A \vdash B$, then $A \vdash' B$. Therefore, if w is a $\vdash'$ saturated theory containing A, it must also contain B. Conversely, if $A \nvdash B$, then $\langle \{A\},\{B\} \rangle$ is $\vdash$ - consistent and therefore $\vdash'$ consistent. By Lemma 1.13, then, there is a $\vdash'$ - saturated set containing A but not B.

Definition 1.34b. If $\Box$ is a connective of L, then $\|\Box\|_{\vdash}$ is the relation which holds between w and $(U_1,\ldots,U_n)$ if and only if there are sentences $A_1,\ldots,A_n$ in L such that $U_1 = \|A_1\|_{\vdash},\ldots, U_n = \|A_n\|_{\vdash}$ and w is a member of $W_{\vdash}$ which contains $\Box A_1\ldots A_n$.

Definition 1.35b. A canonical modal model for $\vdash$ is a modal model $\langle W,0,D,P,\overline{C},a\rangle$ suitable for L such that $w = W_{\vdash}$, and the following hold. D is the set of all individual variables of L'. For each n-ary P in L, $\overline{P} \in P$ is defined by: $\overline{P}(u) = \{(a(x_1),\ldots,a(x_n)) : P(x_1,\ldots,x_n) \in u\}$. Finally, e is defined as in Definition 1.35a.

Lemma 1.12b. If M is a canonical modal model for $\vdash$, then $(M,w) \models A$ if and only if $A \in w$.

Proof. Let $M = \langle W,o,D,P,\overline{C},a\rangle$. We prove the lemma by induction on the length of A.

- If $A = Px_1\ldots x_n$ then $(M,w) \models A$ iff $(a(x_1),\ldots,a(x_n)) \in \overline{P}(w)$ iff $Px_1\ldots x_n \in w$.
- If $A = \Box A_1\ldots A_n$ the proof proceeds exactly as in Lemma 1.12a except that the application of Lemma 1.11a is replaced by an application of Lemma 1.11b.
- If $A = \exists xB$ then $(M,w) \models A$ implies there is a model M' such that $(M',w) \models B$ and $M' = \langle W,0,D,P,\overline{C},a'\rangle$ where, for all variables y not identical to x, $a'(y) = a(y)$. But M' is also a canonical model for $\vdash$, so B is in w by the induction hypothesis.

Since $B \vdash' B$ and $\vdash'$ has classical quantifiers, $B \vdash \exists xB$. Hence $\exists xB$ must be in w. To prove the other direction, suppose $\exists xB(x)$ is in w. Since w is $\vdash'$ saturated, there is a variable x_0 such that $B(x_0)$ is in w. By induction hypothesis, $(M,w) \models B(x_0)$. Therefore $(M,w) \models \exists zB(z)$ where z occurs nowhere in $B(x_0)$ and $B(z)$ is the result of substituting z for x in $B(x)$ (which is not necessarily the same as the result of substituting z for x_0 in $B(x_0)$). Hence by Lemma 1.1(b), $(M,w) = \exists xB(x)$.

• If $A = \forall xB(x)$ the proof is similar to the last case.

In what follows, L is again either a predicate or propositional language.

Theorem 1.5 (Completeness). Let $\vdash$ be a consequence relation on L. Then there is a normal class $\mathcal{M}$ of models suitable for L such that $\vdash$ is strongly complete for $\mathcal{M}$, and the following hold:

(a) If $\vdash$ is partially classical then each member of $\mathcal{M}$ is partially classical.

(b) If $\Box$ is normal Kripke (normal K4, normal S4) with respect to $\vdash$, then for all M in $\mathcal{M}$, $\bar{\Box}_M$ is a point-to-point (transitive point-to-point, transitive and reflexive point-to-point) relation.

(c) If $\vdash$ has classical negation and $\Box$ is normal S5 with respect to $\vdash$, then for all M in $\mathcal{M}$, $\bar{\Box}_M$ is a transitive, reflexive and symmetric point-to-point relation.[12]

(d) If $\vdash$ has classical negation and $(\Box, \boxminus)$ is a pair of tense connectives with respect to $\vdash$ then, for all M in $\mathcal{M}$, $\bar{\Box}_M$ and $\bar{\boxminus}_M$ are transitive and reflexive and $\bar{\Box}_M$ is the converse of $\bar{\boxminus}_M$.

Let $\mathcal{M}^-$ be the set of all canonical modal models suitable for L. First, suppose $\vdash$ is partially classical.

<u>Claim</u>: If '$\wedge$' is a connective of L and if $U = |A|_M$ and $V = |B|_M$ for some A, B in L and some M in $\mathcal{M}^-$, $\bar{\wedge}\, w\, UV$ iff $\langle w,U,V \rangle \in \text{Conj } W$.

<u>Proof</u>. $\bar{\wedge}\, w\, UV$ iff there are C, D in L such that $|A|_M = U = \|C\|_\vdash$ and $|B|_M = V = \|D\|_\vdash$ and $C \wedge D$ is in w. By Lemmas 1.12a, 1.12b $|A|_M = \|A\|_\vdash$ and $|B|_M = \|B\|_\vdash$. By Lemmas 1.11a, 1.11b, $A \dashv\vdash C$ and $B \dashv\vdash D$. Since $\wedge$ is Montague, $A \wedge B \dashv\vdash C \wedge D$. Therefore $\bar{\wedge}\, w\, UV$ iff A B is in w. But by Lemma 1.10 this holds iff A and B are in w, i.e., iff $w \in \|A\|_\vdash \cap \|B\|_\vdash$. This proves the claim. Analagous arguments

---

[12] If R is a binary relation on W then call a point v <u>R-accessible</u> from U if there is a finite sequence $u_1,\ldots,u_n$ of members of W such that $u\, Ru_1,\ldots,u_{n-1}\, Ru_n$ and $u_n\, Rv$. It is easy to check that if $\Box$ is the only non-Boolean connective of L we can delete all points in a model M not $\bar{\Box}$-accessable from the designated point, without affecting the truth in M of any sentence. If we do this for a connective $\Box$ such that $\bar{\Box}$ is an equivalence relation, the remaining points will all be related by $\bar{\Box}$. In view of these remarks (c) explains why in a language with one non-Boolean connective $\Box$ which is S5 with respect to $\vdash$ the truth clause for $\Box$ is often written $(M,w) \models A$ iff for all $w \in W_s$ $(M,w) \models A$. If there are other non-Boolean connectives present, however, this does not hold.

establish that whenever L contains the appropriate connective and U and V are definable in M,

$\bar{\vee}$ w, U, V iff $\langle w,U,V\rangle \in$ Disj W;

$\bar{\rightarrow}$ w, U, V iff $\langle w,U,V\rangle \in$ Impl W; and

$\bar{-}$ w, U iff $\langle w,U\rangle \in$ Neg W.

If $\vdash$ is not partially classical, let $\mathcal{M} = \mathcal{M}^-$. Otherwise, let $\mathcal{M}$ be the class of all models which can be obtained from members of $\mathcal{M}^-$ by replacing $\bar{\wedge}$, $\bar{\vee}$, $\bar{\rightarrow}$, $\bar{-}$ by Conj $W_\vdash$, Disj $W_\vdash$, Impl $W_\vdash$ and Neg $W_\vdash$, respectively.

(i) $\vdash$ is complete for $\mathcal{M}$. By Lemma 1.2 and the previous claim it is sufficient to show $\vdash$ complete for $\mathcal{M}^-$. Suppose $\Gamma \vdash \Delta$ for $\Gamma \subseteq L$, $\Delta \subseteq L$. Then if $M \models \Gamma$ $(M \in \mathcal{M}^-)$ then $\Gamma \subseteq O_M$. Since O_M is $\vdash$ consistent all of Δ can't be in $L - O_M$, and so $\Gamma \models_{\mathcal{M}} . \Delta$. If $\Gamma \nvdash \Delta$ then $\langle \Gamma,\Delta\rangle$ is $\vdash$ consistent. Hence $\langle \Gamma,\Delta\rangle$ can be extended to a $\vdash$ maximal consistent theory w (if L is propositional) or a $\vdash$ saturated theory w (if L is predicate). In either case there is an $M^w \in \mathcal{M}^-$ such that $M^w \models \Gamma$ and for all B in Δ, $M^w \not\models B$. Hence $\Gamma \not\models_{\mathcal{M}} - \Delta$.

(ii) From the definition of $\mathcal{M}$ it is clear that, if $\vdash$ is partially classical, each member of $\mathcal{M}$ is partially classical and, if $\Box$ is Kripke with respect to $\vdash$, that for each M in $\mathcal{M}$ $\bar{\Box}_M$ is point-to-point.

(iii) Suppose $\Box$ is K4 with respect to $\vdash$ and $u \bar{\Box} v$, $v \bar{\Box} w$. We show that $u \bar{\Box} w$. It suffices to show for all A, $\Box A \in u$ implies $A \in w$. But since $\Box A \vdash \Box A$ and $\Box$ is K4, $\Box A \vdash \Box\Box A$. Hence $\Box A \in u$ implies $\Box\Box A \in u$. Since $u \, \|\Box\|_{\vdash} \, v$, then $\Box A \in v$. But $v \, \|\Box\|_{\vdash} \, w$ so $A \in w$.

(iv) Suppose $\Box$ is S4 with respect to $\vdash$. By (iii) $\bar{\Box}$ is transitive; it remains to show that $\bar{\Box}$ is S4. We must show that for all w in W $w \, \|\Box\|_{\vdash} \, w$. But $A \vdash A$ and $\Box$ S4 with respect to $\vdash$, so $\Box A \vdash A$. Hence $\Box A \in w$ implies $A \in w$, and (by the definition of $\|\Box\|_{\vdash}$) $w \, \|\Box\|_{\vdash} \, w$.

(v) Suppose $\vdash$ has classical negation and $\Box$ is S5 with respect to $\vdash$. By (iv) $\bar{\Box}$ is transitive and reflexive. We must show that for all u, v in W $u \, \|\Box\|_{\vdash} \, v$ implies $v \, \|\Box\|_{\vdash} \, u$, i.e., that if v contains all A such that A is in u, then u contains all B such that B is in v. So suppose $u \, \|\Box\|_{\vdash} \, v$ and $B \notin U$. Since $\Box B \vdash \Box B$ and $\vdash$ has classical negation, $\phi \vdash -\Box B, \Box B$. Since $\Box$ is S5, this means $\phi \vdash \Box - \Box B, \Box B$. By classical negation $-\Box B \vdash \Box - \Box B$. But we also know $B \vdash B$, so because $\Box$ is S4, $\Box B \vdash B$ and, because $\vdash$ has classical negation, $-B \vdash -\Box B$. An application of cut yields $-B \vdash \Box - \Box B$. Now since B is not in u. By Lemma 1.12a or b, $-B \in U$. Therefore $\Box - \Box B$ is also in U. But $u \, \|\Box\|_{\vdash} \, v$, so $-\Box B$ is in v, i.e., $\Box B$ is not in v.

(vi) Suppose $\vdash$ has classical negation and $(\Box, \boxminus)$ is a pair of tense connectives. By (iv), $\bar{\Box}$ and $\bar{\boxminus}$ are transitive and reflexive. Claim: For all u, v in W, $u \|\Box\|_{\vdash} v$ iff $v \|\boxminus\|_{\vdash} u$.

Proof: Suppose $u \|\Box\| v$ and $\boxminus A \in v$. By reasoning similar to (v) we get $-A \vdash \Box - \boxminus A$. Hence if A weren't in u then $\Box - \boxminus A$ would be. Since $u \|\Box\|_{\vdash} v$ therefore A would not be in v. Contradiction. The other direction is proved similarly.

We have considered a number of conditions which we described as properties of connectives. For the following corollaries to the completeness theorem we reformulate these as axioms and rules for consequence relationa and logics. For easy reference we collect and label some of them below:

CC: If $\wedge$ is a connective of L, $\Gamma, A, B \vdash \Delta$ iff $\Gamma, A \wedge B \vdash \Delta$;

CD: If $\vee$ is a connective of L, $\Gamma \vdash A, B, \Delta$ iff $\Gamma \vdash A \vee B, \Delta$;

CI: If $\rightarrow$ is a connective of L, $\Gamma, A \vdash B, \Delta$ iff $\Gamma \vdash A \rightarrow B, \Delta$;

CN: If $-$ is a connective of L, $\Gamma, A \vdash \Delta$ iff $\Gamma \vdash -A$ and $\Gamma \vdash A, \Delta$ iff $\Gamma\, -A \vdash \Delta$.

$M_{\Box}$: $\Box$ is n-ary and if $A_i \dashv\vdash B_i$ for $1 \leq i \leq n$, then $\Box A_1 \ldots A_n \dashv\vdash \Box B_1 \ldots B_n$;

$K_{\Box}$: $\Box$ is unary and if $\Gamma \vdash B$, then $\Box\Gamma \vdash \Box B$;

$4_{\Box}$: $\Box$ is unary and if $\Gamma \vdash B$, then $\Box\Gamma \vdash \Box B$;

$R_{\Box}$: $\Box$ is unary and if $A, \psi \vdash H$ then $\Box A, \psi \vdash H$;

$S_{\Box}$: $\Box$ is unary and if $\Box\Gamma \vdash \Box\Delta, B$ then $\Box\Gamma \vdash \Box\Delta, \Box B$;

$T_{\Box\boxminus}$: $\Box$ and $\boxminus$ are unary and if $\boxminus\Gamma \vdash \Box\Delta, B$ then $\boxminus\Gamma \vdash \Box\Delta, \boxminus B$,

if $\Box\Gamma \vdash \boxminus\Delta, B$ then $\Box\Gamma \vdash \boxminus\Delta, \Box B$.

MP : If $A \in L$ and $A \to B \in L$ then $B \in L$;

$N_{\Box}^{L}$: If $A \in L$ then $\Box A \in L$;

$K_{\Box}^{L}$: $\Box(A \to B) \to (\Box A \to \Box B) \in L$;

$4_{\Box}^{L}$: $\Box A \to \Box\Box A \in L$;

$R_{\Box}^{L}$: $\Box A \to A \in L$;

$5_{\Box}^{L}$: $A \to \Box - \Box - A \in L$;

$T_{\Box}^{L}$: $A \to \boxminus - \Box - A \in L$ and $A \to \Box - \boxminus - A \in L$;

D: The following are in L:

$A \to (B \to A)$, $A \to A$, $(A \wedge B) \to (B \wedge A)$,

$(A \vee B) \to (B \vee A)$, $(A \wedge (B \wedge C)) \to ((A \wedge B) \wedge C)$,

$(A \vee (B \vee C)) \to ((A \vee B) \vee C)$, $(A \wedge B) \to A$,

$A \to (B \vee A)$, $A \to (A \wedge A)$, $(B \vee B) \to B$,

$(A \wedge B) \to (C \vee D) \to (A \to ((B \to C) \vee D))$,

$(A \to B) \to ((B \to C) \to (A \to C))$,

$(A \to (B \vee Q)) \to (((Q \wedge C) \to D) \to ((A \wedge C) \to (B \vee D)))$.

CN: The following are in L:

$(-A \to -B) \to (A \to B)$, $(A \vee B) \to (-A \to B)$, $(-A \to B) \to (A \vee B)$,

$(A \wedge B) \to -(A \to -B)$, $-(A \to -B) \to (A \wedge B)$.

Let L be an arbitrary propositional language, all of whose non-Boolean connectives are unary, and let $\vdash^1_{\mathrm{L}}$ ($\vdash^2_{\mathrm{L}}$, $\vdash^3_{\mathrm{L}}$, $\vdash^4_{\mathrm{L}}$) be the smallest consequence relation $\vdash$ on the propositional language L such that: i) $\vdash$ is closed under CC, CD, CI and CN, and ii) for all non-Boolean connectives $\Box$, $\vdash$ is closed under $K_\Box$ ($K_\Box$ and $4_\Box$; $K_\Box$, $4_\Box$, and $R_\Box$, $K_\Box$, $4_\Box$, $R_\Box$ and $S_\Box$). If L is Boolean, let L^1_{L} (L^2_{L} , L^3_{L} , L^4_{L}) be the smallest set $L \subseteq \mathrm{L}$ such that: i) L satisfies D; ii) for all non-Boolean connectives, $\Box$, L satisfies $K^L_\Box$ ($K^L_\Box$ and $4^L_\Box$; $K^L_\Box$, $4^L_\Box$, and $R^L_\Box$; $K^L_\Box$, $4^L_\Box$, $R^L_\Box$ and $S^L_\Box$); and iii) L satisfies MP and $N^L_\Box$ for all non-Boolean $\Box$. Let M^1_{L} (M^2_{L} , M^3_{L} , M^4_{L}) be the class of all partially classical modal models M suitable for L such that for all non-Boolean connectives $\Box$, $\bar{\Box}_{\mathrm{M}}$ is a point-to-point (transitive point-to-point; transitive and reflexive point-to-point; transitive, reflexive and symmetric point-to-point) relation.

<u>Corollary</u> (Axiomatization of Consequence Relations).

$$\models_{M^i_{\mathrm{L}}} = \vdash^i_{\mathrm{L}} \qquad \text{for } i = 1,2,3,4.$$

<u>Proof</u>. By Theorem 1.4, $\vdash^i_{\mathrm{L}} \subseteq \models_{M^i_{\mathrm{L}}}$ for $i = 1,2,3,4$. Conversely, by Theorem 1.5, $\vdash^i_{\mathrm{L}} = \models_M$ for some $M \subseteq M^i_{\mathrm{L}}$ for $i = 1,2,3,4$. But, for all classes M, M' of models, if $M \subseteq M'$ then $\models_{M'} \subseteq \models_M$. Hence $\models_{M^i_{\mathrm{L}}} \subseteq \vdash^i_{\mathrm{L}}$ for $1 \leq i \leq 4$.

<u>Corollary</u> (Axiomatization of Logics).

$\mathcal{L}^i_L$ is complete for $\mathcal{M}^i_L$ (i = 1,2,3,4).

<u>Proof</u>. By Theorems 1.2 and 1.3, $\mathcal{L}^i_L = \mathcal{L}(\vdash^i_L)$. Hence $A \in \mathcal{L}^i_L$ iff $\vdash^i_L A$ iff $\models_{\mathcal{M}^i_L} A$.

<u>Note</u>: If L and L' each contain only a single non-Boolean connective then $\mathcal{L}^1_L$, $\mathcal{L}^2_L$, $\mathcal{L}^3_L$ and $\mathcal{L}^4_L$ are the logics usually labelled K, K4, S4, and S5, respectively. It is clear that we could extract axiomatizations for other logics and consequence relations from Theorem 1.5. Two more examples are worth noting:

(1) SC = the smallest classical logic on a language L with no non-Boolean connectives is complete with respect to the class of all classical propositional models for L (and therefore for the class of all classical algebraic models for L).

(2) Pred C = the smallest classical logic on a predicate language with no non-Boolean connectives except the quantifiers is complete with respect to the class of all predicate models for L.

<u>Frames</u> If M = (W, 0 , C, V) is a model suitable for L, the triple F = (W, 0 , $\bar{C}$) is often called a <u>frame</u> for L. A sentence B of L is <u>true</u> in the frame F if, for all models M such that $W_M = W$, $0_M = 0$ and $\bar{C}_M = \bar{C}$, M $\models$ A. It is clear that completeness with respect to a class of frames (defined in the obvious way) is equivalent to completeness with respect to a valuation-unrestricted class of models. The question of which logics are complete with

respect to classes of frames has received a great deal of attention in the recent literature. As we have just seen, K, K4, S4, S5 and SC are all complete for classes of frames. It follows from Theorem 1.4f that if L is not substitution-closed, it will not be complete for a class of frames. The converse does not hold. Substitution-closed logics have been constructed which are complete for a class of models, but which are not complete for any class of frames.[13]

H. Equivalence, Translation and Definability

Thus far we have identified logics with sets of sentences, and consequence relations with sets of pairs of sets of sentences. Within this framework we would be forced to say, for example, that the classical propositional calculus formulated with only the connectives '$-$' and '$\wedge$' primitive is a different logic than the classical propositional calculus formulated in terms of '$-$', '$\rightarrow$'. This certainly does not square with our ideas about what a logic should be. What is needed is a natural equivalence relation on logics and one on consequence relations which will make it possible to lump together the sets which are really the "same" logic or the "same" consequence relation. In this section we take up this problem, as well as some closely related matters concerning when a consequence relation is a fragment of another, and when a connective is definable from others (with respect to a consequence relation).

---

[13] See [Fine, 1974] and [Thomason, 1974].

All of this will be useful in later chapters. We shall deal explicitly only with consequence relations, but it should be clear that everything we say can be applied equally well to logics which meet appropriate conditions.

Suppose first that we are interested only in the behavior of individual sentences under consequence relations. From the definition of consequence relation, we can see that a consequence relation restricted to singleton sets is just a weak partial order on the set of sentences. We might be tempted to say that $\vdash$ and $\vdash'$ are equivalent if they determine isomorphic weak partial orders, i.e., if there is a 1-1 function f from L onto L' such that $A \vdash B$ if and only if $f(A) \vdash' f(B)$. But this is too strong. For when we study consequence relations we are only interested in the distinctions that can be made with the consequence relation. Whether A and B are different sentences does not matter so much as whether there is some sentence C such that $C \vdash A$ and $C \nvdash B$ or such that $A \vdash C$ and $B \nvdash C$. If $\vdash$ is Montague, then if $A \dashv\vdash B$ there will not be any way to distinguish them. As far as $\vdash$ is concerned, we may as well identify them. Speaking loosely, to require the two weak orders to be isomorphic would be to insist that there be as many ways to express a single idea in L' as there are in L. But we are only interested in which ideas can be expressed, not in how many ways there are to express them. Let us, therefore take our earlier remarks seriously and identify two sentences which bear the relation $\dashv\vdash$ to one and other. Since $\vdash$ is a weak partial order $\dashv\vdash$ is an equivalence relation, and $\vdash$

induces a partial order on the resulting equivalence classes. As a first approximation, then, we take two consequence relations to be "the same" if they determine isomorphic partial orders.

Definition 1.37. $\vdash$ is similar to $\vdash'$ (written '$\vdash \sim \vdash'$') if there is a 1-1 function from $L/\dashv\vdash$ onto $L'/\dashv\vdash'$ such that, for all A, B in L $[A] > [B]$ iff $f([A]) >' f([B])$ (where $[A] > [B]$ means that for all A in [A], and all B in [B], $A \vdash B$).

In practice, it is not so easy to find the function required by this definition. For this reason we give an alternative definition and prove it equivalent to the original one.

Definition 1.38a. $\vdash$ is similar to $\vdash'$ via f_1, f_2 (written '$\vdash \overset{f_1 f_2}{\sim} \vdash'$') if f_1 and f_2 are functions such that $f_1 : L \to L'$, $f_2 : L' \to L$ and the following hold for all A, B in L and all C, D in L':

1) $A \vdash B$ implies $f_1(A) \vdash' f_1(B)$;
2) $C \vdash' D$ implies $f_2(C) \vdash f_2(D)$;
3) $A \dashv\vdash f_2(f_1(A))$;
4) $C \dashv\vdash' f_1(f_2(C))$.

Theorem 1.6. $\vdash \sim \vdash'$ iff there are f_1, f_2 such that $\vdash \overset{f_1 f_2}{\sim} \vdash'$.

Proof. (a) Suppose $\vdash \sim \vdash'$, and let F be the appropriate bijection. Let $C: L/\dashv\vdash \cup L'/\dashv\vdash' \to L \cup L'$ be any function which satisfies the condition $C([A]) \in [A]$ and define $f_1 : L \to L'$ by

$f_1(A) = C(F([A]))$ and $f_2 : L' \to L$ by $f_2(C) = CF^{-1}([C])$. It is then easy to check that f_1 and f_2 satisfy conditions 1-4 of Definition 1.38a.

(b) Suppose $\vdash \overset{f_1 f_2}{\sim} \vdash'$. For all $[A]$ in $L/\dashv\vdash$ and all $[C]$ in $L'/\dashv\vdash'$, let $F([A]) = [f_1(A)]$.

1) F is well defined. $[A] = [B]$ implies $A \dashv\vdash B$ implies $f_1(A) \dashv\vdash f_1(B)$ implies $[f_1(A)] = [f_1(B)]$ implies $F([A]) = F([B])$.

2) F is 1-1. $F([A]) = F([B])$ implies $[f_1(A)] = [f_1(B)]$ implies $f_1(A) \dashv\vdash f_1(B)$ implies $f_2(f_1(A)) \dashv\vdash' f_2(f_1(B))$ implies $A \dashv\vdash B$ implies $[A] = [B]$.

3) F is onto. Let $[C] \in L'/\dashv\vdash'$. $C \dashv\vdash f_1(f_2(C))$ implies $[C] = [f_1(f_2(C))] = F([f_2(C)])$.

4) F is order preserving. $[A] > [B]$ implies $A \vdash B$ implies $f_1(A) \vdash' f_1(B)$ implies $F([A]) >' F([B])$.

5) F^{-1} is order preserving. $F([A]) >' F([B])$ implies $f_1(A) \vdash' f_1(B)$ implies $f_2(f_1(A)) \vdash f_2(f_1(B))$ implies $A \vdash B$ implies $[A] > [B]$.

1) - 5) imply $\vdash \sim \vdash'$.

Definition 1.38a, we said, was introduced to provide a more convenient account of similarity than that provided by Definition 1.37. It also has another advantage, however: it can easily be extended to cover the behavior of sets of sentences under $\vdash$. We will use '$f(\Gamma)$' where Γ is a set of sentences to mean the range of f restricted to Γ, i.e., $f(\Gamma) = \{f(A) : A \in \Gamma\}$. We will also use '$f((A,B))$' to mean $(f(A),f(B))$ and '$f((\Gamma,\Delta))$' to mean $(f(\Gamma),f(\Delta))$.

<u>Definition 1.38b</u>. <u>$\vdash$ is strongly similar to $\vdash'$ via f_1,f_2</u> (written $\vdash \overset{f_1 f_2}{\approx} \vdash'$) if $f_1 : L \to L'$, $f_2 : L' \to L$ and the following conditions hold.

1) For all $\Gamma,\Delta \subseteq L$, $\Gamma \vdash \Delta$ implies $f_1(\Gamma) \vdash' f_1(\Delta)$.

2) For all ψ, $\Theta \subseteq L'$, $\psi \vdash' \Theta$ implies $f_2(\psi) \vdash f_2(\Theta)$.

3) For all A in L, $A \dashv\vdash f_2(f_1(A))$.

4) For all C in L', $C \dashv\vdash' f_1(f_2(C))$, $\vdash$ is <u>strongly similar</u> <u>to $\vdash'$</u> if there are f_1,f_2 such that $\vdash \overset{f_1 f_2}{\approx} \vdash'$.

In general we <u>are</u> interested in the behavior of sets rather than single sentences and therefore we shall not often be satisfied by weak similarity. The next theorem, however, states conditions under which we need not look beyond weak similarity. Before stating it, we prove the following lemma.

<u>Lemma 1.14</u>. Suppose $\vdash$ and $\vdash'$ are deductive and $\vdash \overset{f_1 f_2}{\sim} \vdash'$. Then

a) $f_1(A \wedge B) \dashv\vdash' f_1(A) \wedge f_1(B)$.

b) $f_1(A \vee B) \dashv\vdash' f_1(A) \vee f_1(B)$.

Proof. We prove a). Since $\vdash$ is deductive, $A \wedge B \vdash A$ and $A \wedge B \vdash B$. Hence $f_1(A \wedge B) \vdash' f_1(A)$ and $f_1(A \wedge B) \vdash' f_1(B)$. Therefore $f_1(A \wedge B) \vdash f_1(A) \wedge f_1(B)$. For the other direction, note that $f_1(A) \wedge f_1(B) \vdash' f_1(A)$ so $f_2(f_1(A) \wedge f_1(B)) \vdash f_2(f_1(A)) \vdash A$. Similarly $f_2(f_1(A) \wedge f_1(B)) \vdash B$. Therefore, $f_2(f_1(A) \wedge f_1(B)) \vdash A \wedge B$. This means $f_1\ f_2(f_1(A) \wedge f_1(B)) \vdash' f_1(A \wedge B)$, so $f_1(A) \wedge f_1(B) \vdash' f_1(A \wedge B)$.

Theorem 1.7. If $\vdash$ and $\vdash'$ are regular, finitary and deductive, $\vdash \overset{f_1 f_2}{\sim} \vdash'$ iff $\vdash \overset{f_1 f_2}{\approx} \vdash'$.

Proof. The "if" direction is immediate. To prove the "only if" direction, suppose $\vdash \overset{f_1 f_2}{\sim} \vdash'$. We need only verify clauses a and b of Definition 1.41b. Since these are symmetrical, we do only a. Assume $\Pi \vdash \Sigma$. Since $\vdash$ is finitary there are finite or empty sets $\Pi' \subseteq \Pi$, $\Sigma' \subseteq \Sigma$ such that $\Pi' \vdash \Sigma'$.

(i) Suppose Π' and Σ' are both non-empty, say $\Pi' = \{A_1, \ldots, A_m\}$ and $\Sigma' = \{B_1, \ldots, B_n\}$. Since $\vdash$ is deductive, $(A_1 \wedge \ldots \wedge (A_{m-1} \wedge A_m) \ldots) \vdash (B_1 \vee \ldots \vee (B_{n-1} \vee B_n) \ldots)$. Hence by Lemma 1.14, $(f_1(A_1) \wedge \ldots \wedge (f_1(A_{m-1}) \wedge f_1(A_m)) \ldots) \vdash'$ $(f_1(B_1) \vee \ldots \vee (f_1(B_{n-1}) \vee f_1(B_n)) \ldots)$. Since $\vdash'$ is deductive, $f_1(\Pi') \vdash' f_1(\Sigma')$. Hence by expansion $f_1(\Pi) \vdash f_1(\Sigma)$.

(ii) Suppose Π' is empty and $\Sigma' = \{B_1, \ldots, B_n\}$, then $\phi \vdash (B_1 \vee \ldots \vee (B_{n-1} \vee B_n) \ldots)$. So for all A in L,

$A \vdash (B_1 \vee \ldots \vee (B_{n-1} \vee B_n)\ldots)$. Therefore, for all A in L, $f_1(A) \vdash' (f_1(B_1) \vee \ldots \vee (f_1(B_{n-1}) \vee f_1(B_n))\ldots)$. In particular, this holds when $A = f_2 A'$ for any A' in L'. Since $\vdash'$ is deductive, this means that for all A' in L', $f_1(f_2(A')) \vdash' f_1(\Sigma')$. Since $A' \dashv\vdash' f_1(f_2(A'))$, we get for all A', $A' \vdash' f(\Sigma')$. Since $\vdash'$ is regular this means $\phi \vdash f(\Sigma')$.

(iii) If Σ is empty and $\Pi = \{A_1, \ldots, A_m\}$ the proof is dual to (ii).

(iv) If $\Pi' = \Sigma' = \phi$ then for all A, B in L, $A \vdash B$. Hence, for all A, B in L, $f_1(A) \vdash' f_1(B)$. In particular for all A', B' in L', $f_1(f_2(A')) \vdash' f_1(f_2(B'))$. But $A' \vdash' f_1(f_2(A'))$ and $f_1(f_2(B')) \vdash' B'$. Hence $\phi \vdash' \phi$.

Fragments

We have defined a relation (strong similarity) on consequence relations which is clearly an equivalence relation. This notion appears to capture something of what we have in mind when we say that two consequence relations are really the same. We would now like to give a similar analysis of what it means for one consequence relation to be a <u>fragment</u> of the other (in the sense that the partially classical consequence relation on the language with '$-$' and '$\wedge$' contains the partially classical consequence relation with only '$\rightarrow$'). One natural account would be to say this relation holds between $\vdash$ and $\vdash'$ just in case there is an <u>embedding</u> of $L/\dashv\vdash$ in $L'/\dashv\vdash'$. Let us adopt this account tentatively. We will give evidence later that it is close to what we wanted.

<u>Definition 1.39</u>. $\vdash$ <u>is a fragment of</u> $\vdash'$ (written $\vdash \prec \vdash'$) if there is a 1-1 function from $L/\dashv\vdash$ into $L'/\dashv\vdash'$ such that for all A, B in L, $[A] \prec [B]$ iff $f([A]) \prec' f([B])$.

Two words of warning must be issued here. First, '$\prec$' and '$\subseteq$' are quite distinct. Suppose we form $\vdash'$ by adding pairs of sentences to $\vdash$ and closing under expansion and cut. It <u>might</u> turn out that $\vdash \prec \vdash'$. (This will happen, for example, if the sentences added are from a language disjoint from L.) On the other hand, it might turn out that $\vdash' \prec \vdash$. (This will happen, for example, if we add the pairs $(A, \Box A)$ to $\vdash_{S5}$.) When logicians say that one logic is "stronger" than another, the relation they usually have in mind is '$\subseteq$' rather than '$\prec$'.

Second, $\prec$ does not have all the nice properties we might expect from the shape of its name. It is a trivial fact that if $\vdash \sim \vdash'$ then both $\vdash \prec \vdash'$ and $\vdash' \prec \vdash$. The converse, however, is false. Let R_1 = {(a,b) : a and b are rationals in the closed interval [0,3] and $a < b$ } and let R_2 = {(a,b : a and b are rationals in the open interval (0,3) and $a < b$}. Then R_1 is isomorphic to $R_2 \upharpoonright [1,2]$ and R_2 is isomorphic to $R_1 \upharpoonright (1,2)$. But R_1 and R_2 are not themselves isomorphic, for one has a greatest element and the other does not.[14]

Having issued these warnings, we can now continue with our investigation of '$\prec$'. As we saw in connection with similarity,

---

[14] An interesting example is that the classical and intuitionistic propositional consequence relations are (weak) fragments of each other, but are not (weakly) equivalent. See [Kleene 1950] p. 492.

it is more convenient and more fruitful to deal with sentences than with their equivalence classes.

Lemma 1.15. $\vdash \prec \vdash'$ iff there is a function $f : L \to L'$ such that for all A, B in L, $A \vdash B$ iff $f(A) \vdash' f(B)$.

Proof. First, suppose $\vdash \prec \vdash'$. Let F be the embedding which shows this, and let $f : L \to L'$ be any function which has the property that $f(A) \in F([A])$. Then it is easy to check that $A \vdash B$ iff $f(A) \vdash' f(B)$. Second, suppose f is a function with the property mentioned. For all $[A]$ in $L/\dashv\vdash$ let $F([A]) = [f(A)]$. It is easy to check that F is well defined, F is 1-1, and $[A] < [B]$ iff $F([A]) <' F([B])$.

Definition 1.40. $\vdash$ is strongly a frament of $\vdash'$ (written $\vdash \lll \vdash'$) if there is a function $f : L \to L'$ such that, for all Γ and Δ, $\Gamma \vdash \Delta$ iff $f(\Gamma) \vdash f(\Delta)$.

From this definition we can see that, if $L \subseteq L'$ and $\vdash$ is a consequence relation on L, then any one-one function $f : L \to L'$ determines a consequence relation $\vdash'$ which strongly contains $\vdash$ as a fragment. For we can let $\vdash' = \{(\Gamma',\Delta')$: either there is a $(\Gamma,\Delta) \in \vdash$ such that $f(\Gamma) \subseteq \Gamma'$, $f(\Delta) \subseteq \Delta'$ or $\Gamma' \wedge \Delta' \neq \phi\}$. It is easy to check that $\vdash'$ is a consequence relation on L' and that $\vdash \lll \vdash'$. Let us call such a $\vdash'$ the f-extension of $\vdash$.

We promised earlier that we would produce evidence that our definition of 'fragment' is a plausible one. A more literal-minded account than ours would be to say that $\vdash$ is a fragment of $\vdash'$

if $L \subseteq L'$ and $\vdash' = \vdash \cap (2^L \times 2^L)$. But we do not want our account to depend on "accidental" matters, like language. For example we would like to think of the tense logic K_t formulated with 'G' and 'H' as primitive connectives as containing as a fragment the logic K with 'L' as a primitive connective, and even the logic K with 'L' and 'M' as primitive connectives.[15] A natural way to remove the language-dependence from the literal-minded approach is to use our earlier account of similarity. Let us say $\vdash \lll^* \vdash'$ if there is a consequence relation $\vdash^2$ on L_2 which is strongly similar to $\vdash$ and a consequence relation $\vdash^3$ on L^3 which is strongly similar to $\vdash'$ such that $\vdash^2 = \vdash^3 \cap (2^{L_2} \times 2^{L_2})$. We define $\vdash \prec^* \vdash'$ in the same manner, using similarity in place of strong similarity. We now show that this account captures the same notion of fragment as our previous account.

Lemma 1.16.

a) If $\vdash' \sim \vdash^2$, $\vdash^2 \prec \vdash^3$, $\vdash^3 \sim \vdash^4$ then $\vdash' \prec \vdash^4$.

b) If $\vdash' \approx \vdash^2$, $\vdash^2 \lll \vdash^3$, $\vdash^3 \approx \vdash^4$ then $\vdash' \lll \vdash^4$.

The proof is straightforward.

---

[15] 'M' is the connective such that for all A, $MA \dashv\vdash -L-A$. Note that in our previous discussions of these logics we never revealed the shape of any connective except the Boolean ones. The others were referred to by special symbols of the metalanguage like '$\Box$'.

<u>Theorem 1.8</u>. (i) $\prec^* = \prec$; (ii) $\lll^* = \lll$.

<u>Proof</u>. We prove (ii). Suppose $\vdash' \lll^* \vdash^4$. Then for some $\vdash^2$, $\vdash^3$ on L_2, L_3 $\vdash' \approx \vdash^2$, $\vdash^2 = \vdash^3 \cap 2^{L_2} \times 2^{L_2}$ and $\vdash^3 \approx \vdash^4$. But then we can easily see that $\vdash^2 \lll \vdash^3$. (Just let f be the identity function on L_2.) So by the previous lemma, $\vdash'$ $\vdash^4$. Now suppose $\vdash' \lll \vdash^4$. Let f be the function which shows this and let $\vdash^2 = \vdash^4 \cap 2^R \times 2^R$ where R is the range of f. Let $g : R \to L_1$ be any function such that for all A in R, $f((g(A)) = A$. Then $\vdash' \overset{fg}{\approx} \vdash^2$. Furthermore, $\vdash^4 \overset{i,i}{\approx} \vdash^4$, where i is the identity function. Hence $\vdash' <<^* \vdash^4$.

Translations

In defining $\vdash \overset{f_1 f_2}{\approx} \vdash'$, we did not require the functions f_1, f_2 to satisfy any special conditions. But, when we use this definition to prove two systems equivalent, the functions employed will usually be of a special, simple kind. We can use this fact to obtain some extra information.

So far everything we have said in this section applies to both predicate and propositional languages. In order to maintain this generality we will henceforth use the term 'atoms of L' to refer both to sentence letters of a propositional language and atomic sentences of a predicate language. We assume that we have at our disposal an enumeration $(a_1, a_2, \ldots)$ of the atoms of L.

<u>Definition 1.41</u>. An <u>n-ary sentence schema of L</u> is a sentence σ in L such that all the atoms which occur in σ are among

$a_1,\ldots,a_n$, i.e., they are among the first n atoms of L. An <u>n-ary sentence schema set of L</u> is a set of n-ary sentence schemas; an <u>n-ary theory schema of L</u> is a pair (Σ,Δ) such that Σ and Δ are n-ary sentence schema sets; and an <u>n-ary theory schema set of L</u> is a set of n-ary theory schemas. A <u>sentence schema</u> (<u>sentence schema set</u>, <u>theory schema</u>, <u>theory schema set</u>) is anything that is an n-ary sentence schema (n-ary sentence schema set, n-ary theory schema, n-ary theory schema set) for some n.

<u>Definition 1.42</u>. If σ is an n-ary sentence schema (sentence schema set, theory schema, theory schema set) then <u>$\sigma(A_1,\ldots,A_n)$</u> is the result of replacing all occurrences of $a_1,\ldots,a_n$ in σ by $A_1,\ldots,A_n$ respectively. $\sigma(A_1,\ldots,A_n)$ is called an <u>instance</u> of σ.

It is convenient to keep in mind that sentence schemas are just sentences of a special kind. If σ is an n-ary sentence schema of length 0, it must be an a_i for some $i \leq n$. If it is of length ≥ 1 it must be of the form $\Box\sigma_1\ldots\sigma_n$ where each σ_i is also an n-ary sentence schema.

<u>Definition 1.43</u>. A function $t : L \to L'$ is called a <u>translation of L into L'</u> if, for every n-ary connective $\Box$ of L there is an n-ary sentence schema $(t\Box)$ of L' such that for all $A_1,\ldots,A_n$ in L, $t(\Box A_1\ldots A_n) = (t\Box)(t(A_1),\ldots,t(A_n))$. If, in addition, t restricted to the atoms of L is the identity function, then t is called a <u>simple translation of L into L'</u>.

Lemma 1.17. If ε is a sentence schema (sentence schema set, theory schema, theory schema set) of L and t is a translation of L into L' then

a) $t(\varepsilon)$ is an n-ary sentence schema (sentence schema set, theory schema, theory schema set). Call it $(t\varepsilon)$.

b) For all $A_1,\ldots,A_n$ in L, $t(\varepsilon(A_1,\ldots,A_n)) = (t\varepsilon)(t(A_1),\ldots,t(A_n))$.

c) If t is simple, then for all $A_1,\ldots,A_n$ in L, $t(\varepsilon(A_1,\ldots,A_n)) = (t\varepsilon)(A_1,\ldots,A_n)$.

Proof. a), b), c) can be proved simultaneously for ε a sentence schema by induction on the length of ε. The other cases follow immediately.

Lemma 1.18.

a) If $\vdash$ is Montague and σ is an n-ary sentence schema then for all $A_1,\ldots,A_n$ and $B_1,\ldots,B_n$, if $A_i \dashv\vdash B_i$ for $1 \leq i \leq n$, then $\sigma(A_1,\ldots,A_n) \dashv\vdash \sigma(B_1,\ldots,B_n)$.

b) If, in addition, $\vdash$ is finitary, then if Γ and Δ are n-ary sentence schema sets

$\Psi \cup \Gamma(A_1,\ldots,A_n) \vdash \Delta(A_1,\ldots,A_n) \cup \Theta$ implies

$\Psi \cup \Gamma(B_1,\ldots,B_n) \vdash \Delta(B_1,\ldots,B_n) \cup \Theta$

Proof. a) is proved by straightforward induction on the length of σ. To prove b) note that any finite subset of $\Gamma(A_1,\ldots,A_n)$ can be written $\{\sigma_1(A_1,\ldots,A_n),\ldots, \sigma_m(A_1,\ldots,A_n)\}$ where each σ_i is in Γ. Hence we can apply a and cut a finite number of times to replace each $\sigma_1^{\cdot}(A_1,\ldots,A_n)$ by $\sigma_1^{\cdot}(B_1,\ldots,B_n)$. Similarly, we can replace any finite set of sentences $\tau_j(A_1,\ldots,A_n)$ in $\Delta(A_1,\ldots,A_n)$ by $\tau_j(B_1,\ldots,B_n)$. A final application of expansion gives the desired result.

Definition 1.44. $\vdash$ is equivalent to $\vdash'$ via t_1, t_2 (written $\vdash \overset{t_1 t_2}{\equiv} \vdash$) if $\vdash \underset{L,L'}{\overset{t_1 t_2}{\approx}} \vdash$ and t_1 and t_2 are translations. If $\vdash \overset{t_1 t_2}{\equiv} \vdash'$ and t_1, t_2 are simple translations, then $\vdash$ is simply equivalent to $\vdash'$.

Lemma 1.19. If $\vdash \overset{t_1 t_2}{\equiv} \vdash'$, then if $\vdash$ is Montague, so is $\vdash'$.

Proof. Suppose $\vdash \overset{t_1 t_2}{\equiv} \vdash'$ and $\vdash$ is Montague. Let $\Box$ be an n-ary connective of L' and suppose $A_i \dashv\vdash' B_i$ for $1 \leq i \leq n$. Then $t_2(A_i) \dashv\vdash t_2(B_i)$ for $i \leq i \leq n$. Since $\vdash$ is Montague, the lemma tells us that $(t_2\Box)(t_2(A_1),\ldots,t_2(A_n)) \dashv\vdash (t_2\Box)(t_2(B_1),\ldots,t_2(B_n))$, i.e., that $t_2(\Box A_1\ldots A_n) \dashv\vdash t_2(\Box B_1\ldots B_n)$. Hence $t_1(t_2(\Box A_1\ldots A_n)) \dashv\vdash t_1(t_2(\Box B_1\ldots B_n))$. But since $t_1(t_2(A)) \dashv\vdash' A$ for all A in L, this means $\Box A_1\ldots A_n \dashv\vdash \Box B_1\ldots B_n$.

To vindicate, at least in part, our earlier claim that our notion of equivalence would be a useful one we now show that a particularly simple kind of axiomatization of a consequence relation can be obtained from a similar axiomatization of an equivalent consequence relation.

Definition 1.45. A finitary n-place rule schema for L (abbreviated ^{n}R.S.) is a pair (Σ,τ) such that Σ is a finite n-ary theory schema set in L whose members (the active hypotheses of the rule schema) are all finite and τ is a finite, n-ary theory schema in L (the conclusion of the rule schema). If Σ is a set of theories in L, and $\Gamma,\Delta \subseteq L$, we write $\Sigma + (\Gamma,\Delta)$ for $\{\langle \Gamma \cup \Psi, \Theta \cup \Delta\rangle : \langle \Psi,\Theta\rangle \in \Sigma\}$. If (Σ,τ) is a finitary n-place rule schema for L, then a set X is strongly closed under (Σ,τ) if for all $A_1,\ldots,A_n$ in L, and all $\Gamma \subseteq L$, $\Delta \subseteq L$, $\Sigma(A_1,\ldots,A_n) + (\Gamma,\Delta) \subseteq X$ implies $\{\tau\} + (\Gamma,\Delta) \subseteq X$. $\vdash$ is weakly closed under (Σ,τ) if, for all $A_1,\ldots,A_n$ in L, $\Sigma(A_1,\ldots,A_n) \subseteq \vdash$ implies $\tau \in \vdash$.

Definition 1.46. A consequence relation $\vdash$ is simply axiomatized by A, R, R' if A is a set of finite theory schemas and R and R' are sets of ^{n}R.S.'s such that $\vdash$ is the smallest set containing all instances of members of A which is strongly closed under all members of R and weakly closed under all members of R'.

Lemma 1.20. If $\vdash$ is simply axiomatized by A, R, R' then $\vdash$ is finitary.

Proof. It is sufficient to show that if (Σ,τ) is finitary n-place rule schema and $\vdash$ is strongly (weakly) closed under (Σ,τ) then (Σ,τ) preserves virtual finiteness with respect to $\vdash$. Suppose each member of $\Sigma(A_1,\ldots,A_n) + (\Gamma,\Delta)$ is virtually finite with respect to $\vdash$ (where possibly $\Gamma = \Delta = \phi$). Then for all (Ψ,Θ) in $\Sigma(A_1,\ldots,A_n)$ there are finite or empty sets $\Gamma_{(\Psi,\Theta)} \subseteq \Gamma$ and $\Delta_{(\Psi,\Theta)} \subseteq \Delta$ such that $\Psi' \cup \Gamma_{(\Psi,\Theta)} \vdash \Theta' \cup \Delta_{(\Psi,\Theta)}$ (where Ψ' and Θ' are subsets of Ψ, Θ respectively). If we let $\Gamma' = \cup\{\Gamma_{(\Psi,\Theta)} : (\Psi,\Theta) \in \varepsilon(A_1,\ldots,A_n)\}$ and $\Delta' = \cup\{\Delta_{(\Psi,\Theta)} : (\Psi,\theta) \in \Sigma(A_1,\ldots,A_n)\}$ then we can apply expansion to get $\Psi \cup \Gamma' \vdash \Theta \cup \Delta'$ for each (Ψ,Θ) in $\Sigma(A_1,\ldots,A_n)$. Since Σ is finite and all the members of Σ are finite, both Γ' and Δ' must be finite as well. $\vdash$ is closed under (Σ,τ) so $\{\tau(A_1,\ldots,A_n) + (\Gamma',\Delta')\} \subseteq \vdash$. Hence the sole member of $\{\tau(A_1,\ldots,A_n) + (\Gamma,\Delta)\}$ is virtually finite.

<u>Theorem 1.9</u>. If $\vdash$ is simply axiomatized by A, R, R' and $\vdash \overset{t_1 t_2}{\equiv} \vdash'$ then $\vdash'$ is simply axiomatized by B, S, S' where $B = t_1(A) \cup \{ \langle\{\sigma\},\{\tau\}\rangle : \sigma = t_1(t_2(\tau))$ or $\tau = t_1(t_2(\sigma))$ and S (S') is the set of all $t_1((\Sigma,\tau))$ such that (Σ,τ) is in R (R').

To prove this we need the following lemma:

<u>Lemma 1.21</u>. If $\vdash \overset{t_1 t_2}{\equiv} \vdash'$, $\vdash$ and $\vdash'$ are finitary and $\vdash$ is strongly (weakly) closed under the rule schema (Σ,τ), then $\vdash'$ is strongly (weakly) closed under $t_1((\Sigma,\tau))$.

<u>Proof</u>. Suppose the hypothesis of the lemma holds and $(t_1\Sigma)(A_1,\ldots,A_n) + (\Psi,\Theta) \subseteq \vdash'$. In other words, for any theory

schema (Γ,Δ) in Σ, $\Psi \cup (t_1\Gamma)(A_1,\ldots,A_n) \vdash' (t_1\Delta)(A_1,\ldots,A_n) \cup \Theta$. By Lemma 1.18b this implies $\Psi \cup (t_1\Gamma)(t_1t_2A_1,\ldots,t_1t_2A_n \vdash'$ $(t_1\Delta)(t_1t_2A_1,\ldots,t_1t_2A_n) \cup \Theta$.[16] By Lemma 1.17 this is the same as saying $\Psi \cup t_1\ (\Gamma(t_2A_1,\ldots,t_2A_n)) \vdash' t_1(\Delta(t_2A_1,\ldots,t_2A_n)) \cup \Theta$, Hence $t_2\Psi \cup t_2t_1\ (\Gamma(t_2A_1,\ldots,t_2A_n)) \vdash t_2t_1(\Delta(t_2A_1,\ldots,t_2A_n)) \cup t_2\Theta$. Since $\vdash$ is finitary and $t_2t_1A \dashv\vdash A$ for all A in L, it is clear we can apply a finite number of cuts to show that this condition implies $t_2\Psi \cup \Gamma(t_2A_1,\ldots,t_2A_n) \vdash \Delta(t_2A_1,\ldots,t_2A_n) \cup t_2\Theta$. But (Γ,Δ) is an arbitrary member of ε. Hence $\varepsilon(t_2A_1,\ldots,t_2A_n) + (t_2\Psi,\ t_2\Theta) \subseteq \vdash$. Since $\vdash$ is closed under (ε,τ), this means $\{\tau(t_2A_1,\ldots,t_2A_n)\} + (t_2\Psi,\ t_2\Theta) \subseteq \vdash$. Retracing our steps we get $\{t_1(\tau(t_2A_1,\ldots,t_2A_n))\} + (t_1t_2\Psi,\ t_1t_2\Theta) \subseteq \vdash'$ which implies $(t_1\tau)(t_1t_2A_1,\ldots,t_1t_2A_n) + (t_1t_2\Psi,\ t_1t_2\Theta) \subseteq \vdash'$ which implies $(t_1\tau)(A_1,\ldots,A_n) + (\Psi,\Theta) \subseteq \vdash'$.

<u>Proof of theorem</u>. Let $\vdash^*$ be the consequence relation axiomatized by B, S, S'.

a) $\vdash^* \subseteq \vdash'$. It suffices to show that i) $\vdash'$ contains all instances of B and ii) $\vdash'$ is strongly (weakly) closed under S (S').

(i) By the definition of equivalence we know that for all schemes σ, $\sigma \dashv\vdash' t_1t_2\sigma$. By Lemma 1.18b, $\sigma(t_1t_2A_1,\ldots,t_1t_2A_n) \dashv\vdash'$ $t_1t_2\sigma(t_1t_2A_1,\ldots,t_1t_2A_n)$ for arbitrary $A_1,\ldots,A_n$. By Lemma 1.16b, this means $\sigma(A_1,\ldots,A_n) \dashv\vdash' t_1t_2\sigma(A_1,\ldots,A_n)$.

---

[16]Here and in what follows parentheses are omitted from expressions of the form $t_1(t_2(A))$ and $t(A)$ in order to make reading easier.

It remains only to show that the translations of instances of A are in $\vdash'$. (Γ,Δ) in A implies $\Gamma(t_2(A_1),\ldots,t_2(A_n)) \vdash \Delta(t_2(A_1),\ldots,t_2(A_n))$ for arbitrary $A_1,\ldots,A_n$. So by Lemma 1.17b, $(t_1\Gamma)(t_1t_2A_1,\ldots,t_1t_2A_n) \vdash' (t_1\Delta)(t_1t_2A_1,\ldots,t_1t_2A_n)$, and by Lemma 1.18a, $(t_1\Gamma)(A_1,\ldots,A_n) \vdash' (t_1\Delta)(A_1,\ldots,A_n)$.

b) $\vdash' \subseteq \vdash^*$.

<u>Claim</u>: $(\Gamma,\Delta) \in \vdash$ implies $t_1(\Gamma,\Delta) \in \vdash^*$.

<u>Proof</u>. If (Γ,Δ) is an instance of some schema τ in A, then $t_1(\Gamma,\Delta)$ is an instance of the schema $t_1(\Gamma,\Delta)$ in $t_1(A)$. Now suppose $(\Gamma,\Delta) = (\Psi \cup \Psi', \Theta' \cup \Theta)$ where (Ψ',Θ') is an instance $\tau(A_1,\ldots,A_n)$ of the conclusion of a rule (Σ,τ) in R (R') such that $\Sigma(A_1,\ldots,A_n) + (\Psi,\Theta) \subseteq \vdash$ $(\Sigma(A_1,\ldots,A_n) \subseteq \vdash)$, and suppose that $t_1(\Sigma(A_1,\ldots,A_n)) + (t_1\Psi, t_1\Theta) \subseteq \vdash^*$ $(t_1(\Sigma(A_1\ldots A_n)) \subseteq \vdash)$. By Lemma 1.18b, $(t_1\Sigma)(t_1A_1,\ldots,t_1A_n) + (t_1\Psi, t_1\Theta) \subseteq \vdash^*$ $(t_1\Sigma(t_1A_1,\ldots,t_1A_n) \subseteq \vdash^*)$. But $\vdash^*$ is strongly (weakly) closed under $(t_1(\Sigma), t_1(\tau))$, so $(t_1\tau)(t_1A_1,\ldots,t_1A_n) + (t_1\Psi, t_1\Theta) \in -^*$ $(t_1\tau)(t_1A_1,\ldots,t_1A_n) \in \vdash^*$. Applying lemma again, $t_1(\tau(A_1,\ldots,t_1A_n)) + t_1(\Psi,\Theta) \in \vdash^*$ $(t_1(\tau(A_1,\ldots,A_n)) \in \vdash^*)$. This proves the claim. b) now follows easily. For $\Gamma \vdash' \Delta$ implies $t_2(\Gamma) \vdash t_2(\Delta)$, so by the above, $t_1t_2(\Gamma) \vdash^* t_1t_2(\Delta)$. But $\vdash^*$ is finitary and $A \dashv\vdash^* t_1t_2A$ holds for all formulas A, because (A,t_1t_2A) is an instance of some member of B. Hence $\Gamma \vdash^* \Delta$.

In the next lemma, the preceding theorem is strengthened by weakening the hypothesis that t_1 and t_2 be translations between L and L'.

<u>Definition 1.47</u>. If $\Gamma \subseteq L$, then $\vdash$ <u>admits a Γ-normal form</u> if, with each A in L we can associate a sentence $B \in \Gamma$ such that $A \dashv\vdash B$. We call such a B the <u>Γ-normal form of A</u> and denote it by '$\Gamma(A)$'. If $\Theta \subseteq L$ we write '$\Gamma(\Theta)$' for '$\{\Gamma(A) : A \in \Theta\}$'.

<u>Lemma 1.28</u>. Suppose the following hold:

1) $\vdash$, $\vdash^-$ and $\vdash^+$ are finitary consequence relations.

2) $\vdash^+$ admits a Δ-normal form and $\vdash^- = \vdash^+ \cap 2^\Delta \times 2^\Delta$.

3) $\vdash$ is simply axiomatized by (A,R,R') and $\vdash^-$ by (B,S,S').

4) $\vdash \overset{t_1t_2}{\equiv} \vdash^-$.

Then $\vdash'$ is simply axiomatized by (C,T,T') where

$C = t_1(A) \cup B \cup \{ \langle \{\sigma\},\{\tau\} \rangle : \sigma = t_1t_2\tau \text{ or } \tau = t_1t_2\sigma\}$,

$= t_1(R) \cup S$, $T' = t_1(R') \cup S'$ and $T = t_1(R) \cup S$.

(Notice that when $\vdash^- = \vdash^+$, Lemma 1.28 is the same as Theorem 1.9.)

<u>Proof</u>. By the previous lemma we know that $\vdash^-$, and hence $\vdash^+$, contains all instances of $t_1(A) \cup B$. Now suppose $\Psi \vdash^+ \Theta$. Since $\vdash^+$ is finitary and since it admits a Δ-normal form, $\Delta(\Psi) \vdash^+ \Delta(\Theta)$. Hence $\Delta(\Psi) \vdash^- \Delta(\Theta)$. By the previous theorem there is a 'derivation' of this from $t_1(A)$ using $t_1(R)$ and $t_1(R')$. By applying B, S and S' we can them 'derive' $\Psi \vdash^+ \Theta$.

We are now ready to introduce the notion of definability and to show how it is related to equivalence.

Definition 1.47. If $\Box$ is an n-ary connective in L and σ is an n-ary sentence-schema in L which does not contain any occurrence of $\Box$, then $\Box$ is definable by σ in $\vdash$ if, for all sentences $A_1,\ldots,A_n$ not containing $\Box$, $\Box A_1\ldots A_n \dashv\vdash \sigma(A_1,\ldots,A_n)$.

Definition 1.48. Let $\vec{\Box} = (\boxed{1}, \boxed{2}, \ldots)$ be a (possibly infinite) sequence of symbols not in L, and let $\vec{\sigma} = (\sigma_1,\sigma_2,\ldots)$ be a sequence of sentence schemes in L of the same length as $\vec{\Box}$. Then the $(\vec{\Box}/\vec{\sigma})$ definitional extension of $\vdash$ (written $\vdash + (\vec{\Box}/\vec{\sigma})$) is the smallest Montague consequence relation, $\vdash'$, on the language L' obtained by adding the $\boxed{i}$'s to the connectives of L, such that $\vdash \subseteq \vdash'$ and each $\boxed{i}$ is definable by σ_i in $\vdash'$.

Lemma 1.22. Let $\vec{\Box}$, $\vec{\sigma}$, L, L' be as above and let t be the function from L' to L such that $t(p) = p$ for all sentence letters p, $t(\Box A_1\ldots A_n) = \Box t(A_1)\ldots t(A_n)$ for all n-ary connectives $\Box$ of L and $t(\boxed{i}\, A_1\ldots A_n) = \sigma_i(t(A_1),\ldots,t(A_n))$ for all n-ary connectives $\boxed{i}$ in L' - L. Then if $\vdash$ is finitary, $\vdash + (\vec{\Box}/\vec{\sigma}) = \{\langle\Gamma,\Delta\rangle : t(\Gamma) \vdash t(\Delta)\}$.

Proof. Let $\vdash'$ be the set on the right. It will be convenient to identify $\Box p_1\ldots p_n$ with $\Box$. We prove first that $\vdash'$ has the right properties.

(i) $A \in \Gamma \cap \Delta$ implies $t(A) \in t(\Gamma) \cap t(\Delta)$ implies $t(\Gamma) \vdash t(\Delta)$.

(ii) If $\Gamma \vdash' \Delta$ and $\Psi, \Theta \subseteq L'$, then $t(\Gamma) \vdash t(\Delta)$ and $t(\Psi)$, $t(\Theta) \subseteq L$. Hence $t(\Gamma \cup \Psi) \vdash t(\Delta \cup \Theta)$ so $\Gamma, \Psi \vdash' \Delta, \Theta$.

(iii) If $\Gamma \vdash' \Delta, A$ and $\Psi, A \vdash' \Theta$, then $t(\Gamma) \vdash t(\Delta)$, $t(A)$ and $t(A)$, $t(\Psi) \vdash \Theta$. Therefore, $t(\Gamma \cup \Psi) \vdash t(\Delta \cup \Theta)$ so $\Gamma, \Psi \vdash' \Delta, \Theta$.

(iv) Suppose $A_i \dashv\vdash' B_i$ for $1 < i < n$. Then $t(A_i) \dashv\vdash t(B_i)$ for $1 \leq i \leq n$. Hence, by Lemma 1.18a, $(t\Box)(tA_1, \ldots, tA_n)$ $\dashv\vdash (t\Box)(tB_1, \ldots, tB_n)$, i.e., $t(\Box A_1 \ldots A_n) \dashv\vdash t(\Box B_1 \ldots B_n)$. Therefore $\Box A_1 \ldots A_n \dashv\vdash' \Box B_1 \ldots B_n$.

(v) If $\Gamma \vdash \Delta$, then, since $t(\Gamma) = \Gamma$ and $t(\Delta) = \Delta$, $t(\Gamma) \vdash t(\Delta)$. Hence $\Gamma \vdash' \Delta$.

(vi) Suppose $A_1, \ldots, A_n$ are in L. Then $t(\boxed{i}\, A_1 \ldots A_n) = \sigma_i(tA_1, \ldots, tA_n) = \sigma(A_1, \ldots, A_n)$. Since $\sigma_i(A_1, \ldots, A_n)$ is in L, $t(\sigma_i(A_1, \ldots, A_n)) = \sigma_i(A_1, \ldots, A_n)$ also. Hence $t(\boxed{i}\, A_1 \ldots A_n) \dashv\vdash t(\sigma_i(A_1, \ldots, A_n))$ and so $\boxed{i}\, A_1 \ldots A_n \dashv\vdash' \sigma_i(A_1, \ldots, A_n)$.

(i) - (iii) show that $\vdash'$ is a consequence relation. (iv) shows that it is Montague. (v) shows that it contains $\vdash$ and (vi) shows that it contains the necessary definitions. Clearly $\vdash'$ is finitary. It remains only to show that there is no smaller relation meeting these conditions. So suppose there is; say $\vdash^-$.

<u>Claim</u>: $t(A) \dashv\vdash^- A$.

We prove this by induction on the length of A. If $A = p$ for some sentence letter p then $t(A) = A$ so the claim holds for any consequence relation $\vdash^-$. If $A = \Box A_1 \ldots A_n$ then by induction hypothesis, $A_i \dashv\vdash^- t(A_i)$ for $1 \leq i \leq n$. Since $\vdash^-$ is Montague, $A \dashv\vdash^- \Box t(A_1)\ldots t(A_n)$. If $\Box$ is a connective of L then $t(A) = \Box t(A_1)\ldots t(A_n)$ and we're done. If $\Box = \boxed{i}$, then since each $t(A_i)$ is in L we know $\Box t(A_1)\ldots t(A_n) \dashv\vdash^-$ $\sigma_i(t(A_1),\ldots,t(A_n))$. Applying cut we get $A \dashv\vdash^- \sigma_i(t(A_1),\ldots,t(A_n))$ i.e., $A \dashv\vdash^- t(A)$. This proves the claim.

Now suppose $\Gamma \vdash' \Delta$. Then $t(\Gamma) \vdash t(\Delta)$. Since $\vdash$ is finitary, there are finite or empty sets $\Gamma' \subseteq \Gamma$, $\Delta' \subseteq \Delta$ such that $t(\Gamma') \vdash t(\Delta')$. Since $\vdash \subseteq \vdash^-$, $t(\Gamma') \vdash^- t(\Delta')$. But by the claim $tA \dashv\vdash^- A$ for all A in Γ',Δ'. Applying a finite number of cuts we get $\Gamma' \vdash^- \Delta'$. Expanding, $\Gamma \vdash^- \Delta$. Thus $\vdash' \subseteq \vdash^-$ and the lemma is proved.

<u>Theorem 1.10</u>. If $\vec{\Box}$, $\vec{\sigma}$, L, L' and t are as above and $\vdash$, $\vdash'$ are finitary, $\vdash' = \vdash + (\vec{\sigma}/\vec{\Box})$ iff $\vdash \overset{it}{\equiv} \vdash'$ where i is the identity function on L.

<u>Proof</u>. (a) Suppose $\vdash' = \vdash + (\vec{\sigma}/\vec{\Box})$. We show that $\vdash \overset{it}{\equiv} \vdash'$.

(i) Since $\vdash \subseteq \vdash'$, $\Gamma \vdash \Delta$ clearly implies $i(\Gamma) \vdash' i(\Delta)$.

(ii) $\Gamma \vdash' \Delta$ implies (by the previous lemma) $t(\Gamma) \vdash t(\Delta)$.

(iii) If A is in L, $t(i(A)) = t(A) = A$ so $t(i(A) \dashv\vdash A$.

(iv) If C is in L', $i(t(C)) = t(C)$. Claim: $t(C) \dashv\vdash' C$.

Proof: By the previous lemma it is sufficient to show $t(t(C)) \dashv\vdash t(C)$. But $t(C)$ is in L so $t(t(C)) = t(C)$ which proves the claim.

(b) Suppose $\vdash \overset{it}{\equiv} \vdash'$.

(i) $\Gamma \vdash' \Delta$ implies $t(\Gamma) \vdash t(\Delta)$.

(ii) $t(\Gamma) \vdash t(\Delta)$ implies $i(t(\Gamma) \vdash' i(t(\Delta))$, i.e., $t(\Gamma) \vdash' t(\Delta)$. But since $\vdash \overset{it}{\equiv} \vdash'$, $i(t(C)) \dashv\vdash' C$ for all C in L'. Since $\vdash'$ is finitary and $i(t(C)) = C$ this implies that $\Gamma \vdash' \Delta$. By Lemma 1.22, (i) and (ii) are sufficient to prove that $\vdash' = \vdash + (\vec{\sigma} / \vec{\Box})$.

In what follows we assume that C is the set of all connectives of L, that C' is the set of all connectives of L' and that C and C' are disjoint. Since we can always rename the connectives of a language, this is no real restriction.

<u>Theorem 1.11</u>. If $\vdash$, $\vdash'$ are finitary and substitution-closed, then $\vdash$ is simply equivalent to $\vdash'$ iff $\vdash$ and $\vdash'$ have a common definitional extension.

<u>Proof</u>. The "if" part is a corollary of Theorem 1.10. The "only if" part is proved as follows. Let t_1, t_2 be the sentence-letter preserving translations which make $\vdash$ and $\vdash'$ simply equivalent. Let $\vec{\Box} = (\boxed{1}, \boxed{2}, \ldots)$ be an enumeration of all connectives in C' and let $\vec{\bigcirc} = (①, ②, \ldots)$ be an enumeration of all connectives in C. Let $\vec{\sigma} = ((t_2\ \boxed{1}), (t_2\ \boxed{2}), \ldots)$ and $\vec{\tau} = ((t_1\ ①), (t_1\ ②), \ldots)$. (The notation '$t(\Box)$' is used as it was in Lemma 1.22.) Now we define translations t_1^*, t_2^* from the language $(S, C \cup C')$ into $(S, C \cup C')$. For all sentence letters p, let $t_1^*(p) = t_2^*(p) = p$. If $\Box$ is an n-ary connective

in C' let $(t_1^*\Box) = \Box p_1 \ldots p_n$ and let $(t_2^*\Box) = (t_2\Box)$. If $\bigcirc$ is an n-ary connective in C then let $t_1^*(\bigcirc) = (t_1\bigcirc)$ and let $t_2^*(\bigcirc) = \bigcirc p_1 \ldots p_n$. Finally, let $\vdash^* = \{(\Gamma,\Delta) : t_2^*(\Gamma) \vdash t_2^*(\Delta)\}$. We shall show that

$$\vdash^* = \vdash + (\vec{\Box}/\vec{\sigma}) \quad \text{and} \quad \vdash^* = \vdash' + (\vec{\sigma}/\vec{\tau}).$$

The first equation follows directly from Lemma 1.22. To prove the second we must show $\vdash^* = \{(\Gamma,\Delta) : t_1^*(\Gamma) \vdash' t_1^*(\Delta)\}$ i.e., that $t_1^*(\Gamma) \vdash' t_1^*(\Delta)$ iff $t_2^*(\Gamma) \vdash t_2^*(\Delta)$.

Suppose first that $t_1^*(\Gamma) \vdash' t_1^*(\Delta)$. Since $\vdash \overset{t_1 t_2}{\equiv} \vdash'$, $t_2t_1^*(\Gamma) \vdash t_2t_1^*(\Delta)$. Since $\vdash$ is finitary there are $A_1,\ldots,A_m$ in Γ and $B_1,\ldots,B_n$ in Δ such that $t_2t_1^*A_1,\ldots,t_2t_1^*A_m \vdash t_2t_1^*B_1,\ldots,t_2t_1^*B_n$. To prove $t_2^*(\Gamma) \vdash t_2^*(\Delta)$, then, we need only establish the following.

<u>Claim</u>: For all A, $t_2t_1^*(A) \dashv\vdash t_2^*A$. This is proved by induction on the length of A.

- If $A = p$ for some sentence letter p, then
  $t_2t_1^*(A) = t_2(p) = p = t_1^*(A)$.

- If $A = \Box A_1 \ldots A_n$ for $\Box$ in C' then
  $t_1^*A = \Box t_1^*(A_1)\ldots t_1^*(A_n)$ so
  $t_2t_1^*A = (t_2\Box)(t_2t_1^*A_1,\ldots,t_2t_1^*A_n)$ and
  $t_2^*A = (t_2\Box)(t_2^*A_1,\ldots,t_2^*A_n)$.
  But by induction hypothesis, $t_2t_1^*A_i \dashv\vdash t_2^*A_i$ for $1 \le i \le n$, so Lemma 1.18a tells us that $t_2t_1^*A \dashv\vdash t_2^*A$.

- If $A = \bigcirc A_1 \ldots A_n$ for $\bigcirc$ in C, then

$$t_1^*A = (t_1 O)(t_1^*A_1, \ldots, t_1^*A_n) \quad \text{so}$$

$$t_2t_1^*A = t_2((t_1 O)(t_1^*A_1, \ldots, t_1^*A_n)$$

$$= (t_2(t_1 O))(t_2t_1^*A_1, \ldots, t_2t_1^*A_n).$$

$$t_2^*A = O(t_2^*A_1, \ldots, t_2^*A_n).$$

But $(t_2(t_1(O))) = t_2t_1 \; Op_1 \ldots p_n$ (since t_1, t_2 are simple) $\dashv\vdash Op_1 \ldots p_n$. Substituting $t_2t_1^*A_i$ for p_i we get $t_2t_1^*(A) \dashv\vdash O(t_2t_1^*A_1, \ldots, t_2t_1^*A_n)$. Applying induction hypothesis, $t_2t_1^*(A) \dashv\vdash O(t_2^*A_1, \ldots, t_2^*A_n)$, i.e., $t_2t_1^*A \dashv\vdash A$. This proves the claim, and therefore that $t_1^*(\Gamma) \vdash' t_1^*(\Delta)$ implies $t_2^*(\Gamma) \vdash t_2^*(\Delta)$. The other direction is proved in the same manner.

In this section we have up to now taken consequence relations as objects which are simply given. No mention has been made of classes of models which give rise to these relations. But if such classes are known, it is usually much simpler to deal with the models than with the consequence relations themselves. We show how this may be done below.

Lemma 1.23. Let $\mathcal{M}$ and $\mathcal{M}'$ be classes of models suitable for L and L', respectively; let g_1 and g_2 be functions from $\mathcal{M}$ into $\mathcal{M}'$ and $\mathcal{M}'$ into $\mathcal{M}$, respectively; and let f_1 and f_2 be functions from L into L' and L' into L, respectively, such that, for all M and M' in $\mathcal{M}$ and $\mathcal{M}'$ respectively, and all A and C in L and L', respectively,

(i) $M \models A$ iff $g_1(M) \models f_1(A)$.

(ii) $M' \models C$ iff $g_2(M') \models f_2(C)$.

(iii) $M \models A$ iff $M \models f_2 f_1(A)$.

Then $\models_{\mathcal{M}} \overset{f_1 f_2}{\approx} \models_{\mathcal{M}'}$.

Proof.

(a) Suppose $\Gamma \models_{\mathcal{M}} \Delta$.

Let M' be a member of $\mathcal{M}'$ such that $M' \models f_1(\Gamma)$, i.e., such that $M' \models f_1(A)$ for all A in Γ. Then for all A in Γ, $g_2(M') \models f_2 f_1 A$. Hence by (iii) above, $g_2(M') \models \Gamma$. But $g_2(M') \in \mathcal{M}$ so there is some B in Δ such that $g_2(M') \models B$. Applying (iii) again, $g_2(M') \models f_2 f_1 B$. Therefore by (ii) $M' \models f_1(B)$. This shows $f_1(\Gamma) \models_{\mathcal{M}'} f_1(\Delta)$.

(b) Suppose $\Gamma \models_{\mathcal{M}'} \Delta$ and M is a member of $\mathcal{M}$ such that $M \models f_2(\Gamma)$. Then $g_1(M) \models f_1 f_2 \Gamma$, and, therefore $g_2 g_1 M \models f_2 f_1 f_2 \Gamma$. In other words, for all A in Γ, $g_2 g_1 M \models f_2 f_1 f_2 A$. By (iii) $f_2 f_1 f_2 A \mathrel{{}_{\mathcal{M}}{\dashv}\models_{\mathcal{M}}} f_2 A$, so $g_2 g_1 M \models f_2 \Gamma$. By (ii) $g_1(M) \models \Gamma$. Since $\Gamma \models_{\mathcal{M}'} \Delta$ this means that there is a B in Δ such that $g_1(M) \models B$. By (ii) again, $g_2 g_1(M) \models f_2 B$ and by (iii), $g_2 g_1(M) \models f_2 f_1 f_2 B$. By (ii), $g_1(M) \models f_1 f_2 B$ and by (i), $M \models f_2 B$. This shows $f_2(\Gamma) \models_{\mathcal{M}'} f_2(\Delta)$.

(c) By (iii), $A \mathrel{{}_{\mathcal{M}}{\dashv}\models_{\mathcal{M}}} f_2 f_1(A)$.

(d) Suppose M is in $\mathcal{M}'$ and $M \models A$. Then $g_2(M) \models g_2 A$. So by (iii), $g_2(M) \models f_2 f_1 f_2 A$. By (ii) $M \models f_1 f_2 A$. Hence $A \;{}_{\mathcal{M}}\!\dashv\models_{\mathcal{M}'}\; f_2 f_1 A$.

Corollary. If (i) and (ii) of the lemma hold and for all A in L, $g_2 g_1(M) = M$, then $\models_{\mathcal{M}} \overset{f_1 f_2}{\approx} \models_{\mathcal{M}'}$.

Lemma 1.24. Let L' be the language obtained by adding a set $C = \{\boxed{1}, \boxed{2}, \ldots\}$ of new connectives to the connectives of L. For each i such that $\boxed{i}$ is in C, let σ_i be a schema of L. Let $\mathcal{M}'$ be a class of models suitable for L' and for all M' in $\mathcal{M}'$ let $f_2(M')$ be the model for L obtained by deleting from M' the coordinates which interpret the connectives in C. Suppose that, for any n-ary connective $\boxed{i}$ in L, any M' in $\mathcal{M}'$, any sentences $A_1,\ldots,A_n$ in L, and any w in W, $(M',w) \models \boxed{i} A_1 \ldots A_n$ iff $(f_2(M'),w) \models \sigma_i(A_1,\ldots,A_n)$. Then $\models_{\mathcal{M}'} = \models_{\mathcal{M}} + (\vec{\Box}/\vec{\sigma})$ where $\mathcal{M} = \{f_2(M') : M' \in \mathcal{M}'\}$.

Proof. To simplify notation we assume that $C = \{\boxed{1}\}$ and that $\Box$ is the only connective of L. By Theorem 1.8 it is sufficient to show $\models_{\mathcal{M}} \overset{it}{\equiv} \models_{\mathcal{M}'}$ (where i and t are the translations defined in the statement of that theorem). We can show this by Lemma 1.21. Let f_1 be a function from $\mathcal{M}$ to $\mathcal{M}'$ such that for all $M = (W, o, \bar{\Box}, V)$ in $\mathcal{M}$, $f_1(M) = (W, o, \bar{\Box}, \bar{\boxed{1}}, V)$ for some relation $\bar{\boxed{1}}$. (So $f_2(f_1(M)) = M$)

(i) For all B in L', all M' in $\mathcal{M}'$ and all w in $W_{M'}$, $(f_2(M'),w) \models t(B)$ iff $(M',w) \models B$. This is easily proved by induction on the length of B. There are three cases:

$B = p$, $B = \boxed{1} B_1 \ldots B_n$, and $B = \Box B_1 \ldots B_n$. In the first case the proof is immediate. In the second case we must use the hypothesis of the lemma. Finally, the third case must be divided into two subcases according to whether the interpretation of $\Box$ in M' is a point-to-point or a neighborhood relation. In either subcase the result follows from the induction hypothesis and the fact that $\Box$ has the same interpretation in M' as it does in M.

(ii) For all A in L, $M \models A$ iff $f_1(M) \models i(A)$. This is clearly true since $\boxed{1}$ cannot occur in A and M and $f_1(M)$ are identical except for the interpretation of $\boxed{1}$.

(iii) For all M' in $\mathcal{M}'$, $M' \models A$ iff $M' \models i(t(A))$. Since $i(t(A)) = t(A)$ we need to show $M' \models A$ iff $M' \models t(A)$. By (ii) $M' \models A$ iff $f_1(M') \models A$, and by (i) $f_1(M') \models A$ iff $f_2(f_1(M')) \models t(A)$. But $f_2(f_1(M')) = M'$ so this is the desired result.

To complete this chapter we present one final theorem. The proof of this theorem illustrates some of the ideas presented in this section and the result itself is of some interest. It was first proved by Kit Fine.[17]

[17] The proof was communicated to the author orally. It has apparently never been written up.

Theorem 1.12. Every classical Montague consequence relation $\vdash$ is a fragment of some classical normal-Kripke consequence relation $\vdash'$.

Proof. For notational convenience we assume that L has only the single connective $\Box$, which is n-ary. By Theorem 1.5, we know that $\vdash$ is strongly complete for some normal class $\mathcal{M}$ of models. Furthermore, from the proof of this theorem we can see that all of these models have the same set of points $W_{\vdash}$. Let $M = (W, o, \bar{\Box}, V)$ be a model in $\mathcal{M}$. If $\bar{\Box}$ is a point-to-point relation we can replace it by $\bar{\bar{\Box}} = \{(u,V) : u \bar{\Box} v \Rightarrow v \in V\}$ without changing the truth value of any sentence in M. Thus it is safe to assume that in each model in $\mathcal{M}$, $\Box$ is interpreted by a neighborhood relation. Similarly, if $W_{\vdash}$ is finite we can pick a point w_o from $W_{\vdash}$ and add as many copies of w_o to $W_{\vdash}$ as we like. More precisely, if w_o is in W_{-} and A is a set disjoint from $W_{\vdash}$ then let $\mathcal{M} + w_o(A)$ be the set of all models $(W', o', \bar{\Box}', V')$ such that for some $(W, o, \bar{\Box}, V)$ in $\mathcal{M}$, $W' = W \cup A$, $o' \in W'$, $\bar{\Box}' = \{(u,v) : u \bar{\Box} v$, or $w_o \bar{\Box} v$ and $u \in A$, or $u \bar{\Box} w_o$ and $v \in A\}$, and $V'(p) = V(p)$ if $w_o \notin V(p)$, $V'(p) = V(p) \cup A$ otherwise. Clearly, $\mathcal{M} + w_o(A)$ is normal and $\models_{\mathcal{M}+w_o(A)} = \models_{\mathcal{M}}$. Hence we can assume that the set of points of the models of $\mathcal{M}$ is infinite.

Let L^+ be the language obtained by adding to the connectives of L the new unary connectives $\bigcirc$, $\boxed{1}$, ..., $\boxed{n}$, ①, ..., ⓝ. Let L' be the language obtained by removing from the connectives of L^+ the original connective $\Box$ of L. Now for each model $M = (W, o, \bar{\Box}, V)$ of $\Box$ we construct a model

$M^+ = (W, o, \bar{O}, \bar{\boxed{1}}, \ldots, \bar{\boxed{n}}, \bar{①}, \ldots, \bar{ⓝ}, \bar{\Box}, V)$ suitable for L^+ as follows. Let $S^1, S^2, \ldots$ be an enumeration of all n-tuples of sentences in L and let $U = \{u_1, u_2, \ldots\}$ be a countable subset of W. For $1 \leq j \leq n$ let $\bar{\boxed{j}}$ be the relation which holds between u and v iff there is some i such that $u = u_i$ and v is a member of $|A|_M$ where A is the j'th member of the n-tuple S^i. If $V_1, \ldots, V_n$ are subsets of W, we say u <u>codes</u> $(V_1, \ldots, V_n)$ if, for all x in W, x is in V_j iff $u \bar{\boxed{j}} x$. Clearly each u codes a unique n-tuple and each u_i codes $(|(S^i)_1|_M, \ldots, |(S^i)_n|_M)$. Let $u \bar{O} v$ iff there is some n-tuple $A_1, \ldots, A_n$ of sentences of L such that $(M,u) \models \Box A_1 \ldots A_n$ and v codes $(|A_1|_M, \ldots, |A_n|_M)$. Finally, for $1 \leq j \leq n$, let $u \bar{ⓙ} v$ iff it is not the case that $u \bar{\boxed{j}} v$. This completes the definition of M^+.

Let $\mathcal{M}^+ = \{M^+ : M \in \mathcal{M}\}$ and, for all M^+ in $\mathcal{M}^+$ let $f_2(M^+)$ be the model obtained by deleting the penultimate coordinate, $\bar{\Box}$, from M^+. Let $\sigma = -O- (\boxed{1}\, p_1 \wedge \ldots \wedge \boxed{n}\, p_n \wedge ① - p_1 \wedge \ldots \wedge ⓝ - p_n)$.

<u>Claim</u>: For all $A_1, \ldots, A_n$ in L', all M^+ in $\mathcal{M}^+$ and all points w of M^+, $(M^+, w) \models \Box A_1 \ldots A_n$ iff $(f_2(M^+), w) \models \sigma(A_1, \ldots, A_n)$.

<u>Proof</u>. Let $f_2(M^+) = M'$. $(M', w) \models \sigma(A_1, \ldots, A_n)$ iff $\exists v$ such that for some sentences $B_1, \ldots, B_n$ of L, $(M,w) \models \Box B_1 \ldots B_n$, v codes $(|B_1|_M, \ldots, |B_n|_M)$ and $(M', v) \models \boxed{1}\, A_1 \wedge \ldots \wedge \boxed{n}\, A_n \wedge ① -A_1 \wedge \ldots \wedge ⓝ -A_n$. This last clause just says that for all points u of M', $v \bar{\boxed{j}} u$ implies $(M', u) \models A_j$, and not $v \bar{\boxed{j}} u$ implies $(M', u) \not\models A_j$. In other words it says that v

codes $(|A_1|_{M'},\ldots,|A_n|_{M'})$. This means that for $1 \leq j \leq n$ we must have $|A_j|_{M'} = |B_j|_M$. Putting all this together, $(M',w) \models \sigma(A_1,\ldots,A_n)$ iff there are sentences $B_1,\ldots,B_n$ of L such that $(M,w) \models \Box B_1\ldots B_n$ and $B_{i\ M} = |A_i|_{M'}$ for $1 \leq i \leq n$. Since each A_i contains no occurrences of $\Box$, it is clear that $|A_i|_{M'} = |A_i|_M$. Therefore $(M',w) \models \sigma(A_1,\ldots,A_n)$ iff $(M,w) \models \Box A_1\ldots A_n$. This proves the claim. By Lemma 1.22, the claim shows that $\models_{\mathcal{M}^+} = \models_{\mathcal{M}'} + (\Box/\sigma)$. By Theorem 1.4, $\models_{\mathcal{M}'}$ is normal-Kripke and by Theorem 1.8 it is equivalent to $\models_{\mathcal{M}'} + (\Box/\sigma)$. The theorem will follow if we can show that $\models_{\mathcal{M}^+} \cap (2^L \times 2^L) = \models_{\mathcal{M}}$. But if A is in L, then A contains no occurrences of any connective but $\Box$, so for all M in $\mathcal{M}$, $M \models A$ iff $M^+ \models A$. Since $\mathcal{M}^+ \models \{M^+ : M \in \mathcal{M}\}$ the required equation clearly holds.

Remarks on the literature

Much of what was included in Part I is standard material either in elementary logic texts or in standard works on modal logic,[18] but it may be worth pointing out a few things which are not. The material on consequence relations is mostly taken from [Gabbay, 1974], [Scott, 1971], and [Scott, 1975]. The completeness theorem is more general than those given in the literature in two respects. First, the language may have any number of non-classical

[18] e.g., [Hughes, Cresswell 1968], [Segerberg 1971], [Gabbay, f].

connectives. This is possible because the properties of being Kripke, S4, etc., are ascribed to connectives rather than logics. Second, no assumption that there is an underlying classical logic is made. This is possible because the properties of being Kripke, S4, etc., are defined in terms of consequence relations alone. (Most of the definitions are adaptations of the Genteen-style rules given in [Ohnishi, Matsumoto 1961].) All of this is a natural extension of the approach in [Gabbay 1974]. It is worth noting that completeness theorems for fragments of the standard modal logics follow immediately from our general completeness theorem. These have been obtained previously by more tedious methods. (See, for example, [Scharle 1975].) The section on equivalence, translation, and definability is based on [Kotas, Pieczkowski 1970] and [Kanger 1968]. Here again, use of the consequence relation makes it possible for us to say a great deal without the usual presuppositions about the language. After this chapter was mostly written I learned of Kit Fine's unpublished work on the same subject. Fine's notion of equivalence is basically the same as that given here, but his notation is simpler in that a single binary relation $S \subseteq L_1 \times L_2$ ('synonomy') does the work of our functions $f_1 : L_1 \to L_2$ and $f_2 : L_2 \to L_1$.

CHAPTER II

THE MANY-SORTED FRAMEWORKS

A. Introduction

In Chapter I we said that a predicate language contained a single stock of individual variables. Each variable could be interpreted as any member of the domain of individuals of a structure. For some purposes, however, it turns out to be more convenient to view the world as containing several different kinds of individuals, and to include in the language several different kinds of variables, each of which can be interpreted as only one kind of individual. Systems like these have been called <u>many-sorted</u>. They arise naturally in connection with reasoning in certain parts of mathematics. A language in which facts of geometry can be represented, for example, might include different sets of variables for points, lines and planes. Similarly, in the case of linear algebra, different variables for vectors and scalars would be appropriate. It is always possible to reduce these many-sorted frameworks to single-sorted ones by combining the domains of a many-sorted structure and introducing into the language new unary predicate letters whoe interpretations are just the old domains.[1] But many-sorted structures are more natural than their

---

[1] See [Wang 1952].

one-sorted counterparts and many-sorted sentences are more transparent. For we think of geometry and linear algebra as being about several different kinds of objects, not as being about one kind of object which may have one of several properties.

The theme of the next three chapters will be that modal propositional frameworks, like classical predicate frameworks, are often best viewed as 'many-sorted'. Just as variables of different sorts are introduced to range over different kinds of individuals, so sentence letters of different sorts will be introduced to range over different kinds of sentences. There are at least two circumstances under which we shall find it appropriate to distinguish among different kinds of sentences. The first is the case in which our framework is intended to deal with sentences whose truth in a situation depends on different features of the situation. For example, if we wish to represent sentences beginning "It will always be the case that,..." as well as sentences beginning "As far as the eye can see it is the case..." it would be natural to include one kind of sentence letter to be evaluated at times and a second to be evaluated at places.

The second case involves sentences whose truth depends in different ways on the same features of the situation. The natural way to insure that these differences are captured is to place different kinds of restrictions on the interpretations of different sentence letters. In Chapter III, for example, we consider a framework in which the truth of certain sentences depends on what interval of time is associated with them. Some of these sentences (e.g., "John swims in the Channel") have the property that their being true when associated with one interval entails their being true when associated with any subinterval of that interval. Others (e.g., "John swims across the Channel") have the property that their being true when associated with an interval entails their being false at any subinterval of that interval. Clearly the sentence letters which are to represent these sentences must be divided into two groups, and we must place different restrictions on the interpretation of each.

In the remainder of this chapter we consider only sorts corresponding to dependence on different features of the situation. The treatment of sorts corresponding to different kinds of dependence on the same features in unproblematic, and will be illustrated in later chapters. In sections B, C, and D of this chapter we shall introduce the notion of a many-sorted framework and argue for its usefulness. Our initial formulation is suggested by the standard formulation of many-sorted predicate logic, but we shall show that there is a more convenient formulation which will do just as well. In section E we prove some general completeness theorems for many-sorted frameworks. In section F we present a reduction to one-sorted frameworks analagous to the reduction of many-sorted classical predicate systems. In section G we consider the possibility of introducing sort restrictions to block iterations of the necessity operator in S5. The next two chapters deal with more serious applications.

B. Languages

Definition 2.1. If Q is a set, then a Q-sorted language is a pair $L = (S,C)$ where S is a pairwise disjoint collection of countable or empty sets indexed by Q and C is a pairwise disjoint collection of sets indexed by some set of partial functions from finite or empty sequences of members of Q to members of Q. These functions are called types. A type t defined only on the empty sequence $\langle\ \rangle$ is identified with $t(\langle\ \rangle)$. We call the members of Q the sorts of L; the members of each S_j in S, the sort-j sentence letters of L; and the members of each C_r in C, the type-r connectives of L. In practice we usually take the sorts to be either natural numbers of sets of natural numbers and we always take the types to include in their domains of definition only sequences of the same length. If $\Box$ is a type-r connective and the domain of r contains only n-tuples then we say $\Box$ is n-ary.

Definition 2.2. If $L = (S,C)$ is a Q-sorted language, then we define what it means for a finite sequence A to be a sort-j sentence of L by induction on the length of the sequence A.

- If A is of length 1 then A is a sort-j sentence of L iff either $A = (p)$ where $j \in Q$ and $p \in S_j$ or $A = (\Box)$ where $\Box \in C_j$ and $j \in Q$.
- If A is of length $n+1$ then A is a sort-j sentence of L iff $j \in Q$ and $A = (\Box, A_1, \ldots, A_n)$ where $\Box$ is a type-r

connective of L, and for some (not necessarily distinct) $j_1,\ldots,j_n$ in Q, $A_1,\ldots,A_n$ are sentences of sort $j_1,\ldots,j_n$ respectively, and $r(j_1,\ldots,j_n) = j$.

A <u>sentence of L</u> is a sort-j sentence of L for some j in Q. We follow the same conventions in writing sentences of many-sorted languages as we did in writing sentences of one-sorted languages. It is clear that no sentence is a sentence of more than one sort. We can therefore write <u>s(A)</u> for the unique j in Q such that A is a sentence of sort j. Also, if $K \subseteq L$, we write $\underline{K^j}$ for the set of sort-j sentences of K.

Notice that an n-ary connective followed by n sentences is not always a sentence. Suppose, for example, that $L = (S,C)$ is a $\{1,2\}$-sorted language, and $\Box$ is a unary connective of type $r = \{((1),2)\}$. Then there can be no sentences of the form $\Box\Box A$ in L. For any sentence of the form $\Box A$ must be of sort-2 and connectives of type r cannot be applied to sort-2 sentences. The fact that connectives of a many-sorted language cannot always be iterated is noteworthy. The lack of a suitable interpretation for iterated connectives has often proved embarassing to philosophers whose systems seem otherwise to be faithful formalizations of philosophical concepts. In fact the existence of these difficulties is one sign that sort distinctions may be needed. For if a modal connective can be sensibly applied only once to a sentence, then there must be some difference between sentences with that connective and sentences without it.

Henceforth we assume that $L = (S,C)$ is a Q-sorted language.

C. Models -- first formulation

The account of many-sorted propositional languages given above is not unlike the standard account of many-sorted predicate languages. Their interpretation, however, requires a little more thought. Perhaps the most natural way to proceed (since it preserves the analogy with the standard predicate case) is the following.

<u>Definition 2.3a</u>. A <u>model suitable for L</u> is a 4-tuple $(W, o, \bar{C}, V)$ satisfying the following:

(1) W is a pairwise disjoint collection of non-empty sets indexed by the members of Q.

(2) o is a choice set for W, i.e., a set containing exactly one member from each W_j in W.

(3) $\bar{C}$ is a set such that for each type-r connective $\Box$ of L and each $\vec{j} = (j_1,\ldots,j_n)$ in the domain of r, $\bar{C}$ contains a relation $\bar{\Box}(\vec{j})$ such that either $\bar{\Box}(\vec{j}) \subseteq W_{r(\tau)} \times 2^{W_{j_1}} \times \ldots \times 2^{W_{j_n}}$ or $\vec{j} = (j_1)$ and $\bar{\Box}(\vec{j}) \subseteq W_{r(\vec{j})} \times W_{j_1}$.

(4) V is a function such that, for every j in Q, if p^j is a sentence letter of sort-j then $V(p^j) \subseteq W_j$.

The members of each W_j are called <u>j-points</u>. The j-point in o is the <u>designated j-point</u>. o itself is the <u>designated set</u>. Intuitively, each j-point represents that feature of a situation on which the truth of a sort-j sentence depends. The designated j-point represents the appropriate feature of the situation which actually obtains.

We can now define <u>truth of a sort-j sentence at a j-point</u> in much the same way as we defined truth of an ordinary sentence at a point.

<u>Definition 2.4</u>. If $M = (W, o, \overline{C}, V)$ is a model for L, $w_j \in W_j \in W$ and A^j is in L^j then <u>$(M,w^j) \models A^j$</u> iff one of the following holds:

(1) A^j is a sentence letter and $w^j \in V(A^j)$.

(2) A^j is a o-ary connective $\Box$ and $w^j \in \overline{\Box}(\langle\ \rangle)$.

(3) $A^j = \Box A_1 \ldots A_n$, $\overline{\Box}(\vec{j}) \subseteq W_j \times 2^{W_{j_1}} \times \ldots \times 2^{W_{j_n}}$ where $j_i = s(A_i)$ for $1 \leq i \leq n$ and $s(\Box A_1 \ldots A_n) = j$, and finally $\overline{\Box}(\vec{j})\ w, |A_1|_M, \ldots, |A_n|_M$. $(|B|_M = \{u \in W_{s(B)} : (M,u) \models B\}.)$

(4) $A^j = \Box B$, $\overline{\Box}(\vec{j}) \subseteq W_j \times W_{j_1}$ where $j_1 = s(B)$ and $s(\Box B) = j$ and, for all u in W_{j_1}, $(w,U) \in \overline{\Box}(\vec{j})$ implies $(M,u) \models B$.

Truth of a j-sentence in a model is just truth at the designated j-point, i.e., <u>$M \models A^j$</u> iff there is a w_j such that $w_j \in o \cap W_j$ and $(M,w) \models A^j$.

If $\mathcal{M}$ is a class of models suitable for L then we can define <u>validity</u> and <u>consistency</u> of a set $\Gamma \subseteq L$ relative to $\mathcal{M}$ in terms of the truth of sentences of Γ in models of $\mathcal{M}$ exactly as before. (This is true even if Γ contains mixed sorts.)

In Chapter I we argued that if logical truth and logical consequences are to be determined by a class of models, this class of models should be normal, i.e., it should be closed under the operation of changing the designated point of a model. This is no longer the case. For it may turn out that different features of a situation are dependent on one another, so that, say, the j'th feature of the situation's being as represented by w_j precludes the possibility of the k'th feature's being as represented by w_k. For example, suppose the truth of j-sentences depends only on the temperature inside a certain container of gas and the truth of k-sentences depends only on the pressure on the walls of this container. (Plausible candidates for j and k sentences might be "An accurate thermometer would read 70°F," and "An accurate pressure gauge would read 70 lbs/in^2.") We would not want to say that a mixed set Γ of j-sentences and k-sentences was <u>consistent</u> just because there was some temperature at which all the j-sentences were true and some pressure at which all the k-sentences were true; for it might be that this temperature and pressure cannot simultaneously obtain in the container.

The choice of example here may seem unfortunate since it leaves the escape route of distinguishing between logical consistency and "physical consistency". We shall later come across examples where the connections among the sorts are clearly logical. (One such is obtained by making j-sentences temperature-and-pressure dependent and k-sentences merely temperature dependent.) In general, though, we wish to leave the question of what connections among sorts are "logical" to be determined by the framework. Therefore we proceed as follows. For every model we single out a class of compatible choice sets. (To call a choice set "compatible" indicates that all the features represented could obtain simultaneously in some situation.) We then insist that the designated set be a compatible choice set. Finally, for the same reason we restricted our attention to normal classes of models in Chapter I, it is now reasonable to restrict our attention to classes of models closed under the operation of changing the designated point from one compatible choice set to another. Since each j-point is intended to represent a feature of at least one situation, this restriction implies that the classes of models we consider have the following property. If $M = (W, o, \bar{C}, V)$ is in $\mathcal{M}$ and w_j is a j-point of M then there will be at least one model $(W, w, \bar{C}, V)$ such that $w_j \in w$. For want of a better name such classes will be called semi-normal.

D. Models -- second formulation

The account of many-sorted models sketched in the preceding section is different enough from our earlier account of ordinary modal models to make the application of the results and techniques of Chapter I a little difficult. In this section we give an alternative account by which the many-sorted model is seen to be a special kind of ordinary modal model. The idea is that instead of starting with different j-points and then singling out certain sets of these as compatible, we start with points corresponding to compatible choice sets and let equivalence classes of these objects do the work of j-points. Since all of the points represent compatible choice sets we can reasonably require classes of models to be normal. The details are presented below.

Definition 2.3b. A model suitable for L is a 5-tuple, $M = (W, E, o, \overline{C}, V)$ such that

1) W is a non-empty set (the points of M).

2) E is a collection of equivalence relations indexed by the members of Q. If $w \in W$, $S \subseteq W$, and $j \in Q$, we write $[w]_j$ for $\{u \in W\colon u \ E_j \ w\}$ and S/j for $\{[u]_j : u \in S\}$.

3) $o \in W$ (o is the designated point of M).

4) V is a function from sentence letters of L to subsets of W such that if p^j is a sort-j sentence letter and $u \ E_j \ v$ then $u \in V(p^j)$ iff $v \in V(p^j)$.

5) $\overline{C}$ contains a neighborhood or point-to-point relation $\overline{\Box}(\vec{j})$ on W for each type r connective $\Box$ and each $\vec{j}$ in the domain of r such that the following hold. If is type-r, $r((j_1)) = j_o$ and $\overline{\Box}((j_1))$ is point-to-point then $u \ E_{j_o} \ u'$ and $v \ E_{j_1} \ v'$ implies that $u \overline{\Box}((j_1)) v$ iff $u' \overline{\Box}((j_1)) v'$. If $\Box$ is type-r and $r(\vec{j}) = j_o$ where $\vec{j} = (j_1,\ldots,j_n)$, then if $u \ E_{j_o} \ u'$ and for $1 \le i \le n$, $S'/j_i = S/j_i$ then $\overline{\Box}(\vec{j})\ u, S_1,\ldots,S_n$ iff $\overline{\Box}(\vec{j})\ u', S_1',\ldots,S_n'$.

Definitions of truth at a point, truth in a model, and validity, consistency, and consequence with respect to a class of models can be given in exactly the same way as in Chapter I. (The new coordinate E plays no role here.) Points are now regarded as situations (as they were in Part I) and the equivalence relations E_j can be thought of as holding between pairs of situations which agree on their j'th feature. The restrictions placed on the interpretation of sentence letters and connectives guarantee that the truth of a j-sentence depends only on the j'th feature. More precisely, a simple induction will establish the following result.

Lemma 2.1. If $M = (W, E, o, \overline{C}, V)$ is a model, then if A is of sort-j and $w\ E_j\ w'$ then $(M,w) \models A$ iff $(M,w') \models A$.

It is not surprising that the models of this section are not essentially different from those of the preceding one. We make this precise in the next few pages. To avoid confusion we refer to the models introduced in this section as 'D-models' and those of the preceding section as 'C-models'. In subsequent sections this terminology will be dropped and the term 'model' will always refer to D-models.

Lemma 2.2. For every semi-normal collection $\mathcal{M}$ of C-models, there is a normal collection $\mathcal{M}'$ of D-models such that $\models_{\mathcal{M}} = \models_{\mathcal{M}'}$.

Before proving this it is convenient to introduce some notation.

Definition 2.5. Let $M = (W, o, \bar{C}, \bar{V})$ be a C-model suitable for L, and let j be a sort of L. If w is a choice set of W, then the j-part of w (written w^j) is the unique member of $w \cap W_j$. If S is a family of choice sets of W then the j-part of S (written S^j) is the set of all w^j such that w is a member of S.

Proof of Lemma. Let $M = (W, o, \bar{C}, V)$ be a C-model in $\mathcal{M}$, where $W = \{W_j : j \in Q\}$, $o = \{w_j : j \in Q\}$, $\bar{C} = \{\bar{\Box}(\vec{j}) : \Box$ is a type-r connective of L and $\vec{j}$ is in the domain of $r\}$. Now let W' be the set of all choice sets w of W such that $(W, w, \bar{C}, V)$ is in $\mathcal{M}$. For all j in Q, let $E_j = \{(u,v) \in W' \times W' : u^j = v^j\}$. Let $E = \{E_j : j \in Q\}$. For all type-r connectives and all $\vec{j} = (j_1,\ldots,j_n)$ in the domain of r we define a relation $\bar{\bar{\Box}}(\vec{j})$ from $\bar{\Box}(\vec{j})$ as follows. If $\bar{\Box}(\vec{j}) \subseteq W_{r(\vec{j})} \times 2^{W_{j_1}} \times \ldots \times 2^{W_{j_n}}$, then $\bar{\bar{\Box}}(\vec{j}) = \{(w, S_1,\ldots,S_n) : (w^{r(\vec{j})}, S_1^{j_1}, \ldots, S_n^{j_n}) \in \bar{\Box}(\vec{j})\}$. If $\bar{\Box}(\vec{j}) \subseteq W_{r(\vec{j})} \times W_{j_1}$, then $\bar{\bar{\Box}}(\vec{j}) = \{(u,v) : (u^{r(\vec{j})}, v^{j_1}) \in \bar{\Box}(\vec{j})\}$. Let $\bar{C}' = \{\bar{\bar{\Box}}(\vec{j}) : \bar{\Box}(\vec{j}) \in \bar{C}\}$. Finally, let $o' = o$ and let $V'(p^j) = \{w : w^j \in V(P^j)\}$ for all sort-j sentence letters p. It is routine to verify that $M' = (W', E, o', \bar{C}', V')$ is a D-model.

Claim: For all sort-j sentences A^j, $(M',w) \models A^j$ iff $(M,w^j) \models A^j$.

Proof. By induction on the length of A^j. We do the ugliest case: $A^j = \Box B_1 \ldots B_n$ where each B_i is sort j_i and

$\bar{\bar{\Box}}((j_1,\ldots,j_n)) \subseteq W' \times 2^{W'} \times \ldots \times 2^{W'}$. $(M',w) \models A_j$ iff $\bar{\bar{\Box}}(j_1,\ldots,j_n)\ w,\ |B_1|_{M'},\ldots,|B_n|_{M'}$, i.e., iff $\bar{\Box}((j_1,\ldots,j_n))\ w^j,\ (|B_1|_{M'})^{j_1},\ldots,\ (|B_n|_{M'})^{j_n}$. But by induction hypothesis, if B is of sort-j then for all w in W', $w \in |B|_{M'}$ iff $w^j \in |B|_M$. Since M is semi-normal, $\{ w^j : w \in W'\} = W_j$. Hence $(|B|_{M'})^j = |B|_M$. Substituting equals, we get $\bar{\Box}((j_1,\ldots,j_n))\ w^j,\ |B_1|_M,\ldots,|B_n|_M$, iff $(M,w^j) \models A^j$. This proves the claim.

Since $o' = o$ we get immediately that $M \models A$ iff $M' \models A$ for all sentences A. Now let $\mathcal{M}' = \{M' : M \in \mathcal{M}\}$. It is clear that $\models_{\mathcal{M}'} = \models_{\mathcal{M}}$. We must how show that $\mathcal{M}'$ is normal. $M = (W, E, o, \bar{C}, V)$ is in $\mathcal{M}'$ and $w \in W$. Then there is a model $M^- = (W^-, o^-, \bar{C}^-, V^-)$ in $\mathcal{M}$ such that $M = (M^-)'$. Since $w \in W$, there must be a model $(M^-)^w = (W^-, w, \bar{C}^-, V^-)$ in $\mathcal{M}$. From our construction it is clear that $((M^-)^w)' = (W, E, w, \bar{C}, V)$. Hence $\mathcal{M}'$ is normal.

Lemma 2.3. For every normal collection $\mathcal{M}$ of D-models suitable for L there is a semi-normal collection $\mathcal{M}'$ of C-models for L such that $\models_{\mathcal{M}} = \models_{\mathcal{M}'}$.

Proof. Let $M = (W, E, o, \bar{C}, V)$ be a model in $\mathcal{M}$ where $E = \{E_j : j \in Q\}$ and $\bar{C} = \{ \bar{\Box}(\vec{j}) : \Box$ is a type-r connective of L and $\vec{j}$ is in the domain of r$\}$. Let $W' = \{W/j : j \in Q\}$ and let $o' = \{[o]_j : j \in Q\}$. For all type-r connectives $\Box$ of L and all $\vec{j} = (j_1,\ldots,j_n)$ in the domain of r, we define $\bar{\bar{\Box}}(\vec{j})$ from $\bar{\Box}(\vec{j})$ as follows. If $\bar{\Box}(\vec{j}) \subseteq W \times W$ then $\bar{\bar{\Box}}(\vec{j}) = \{([u]_{r(\vec{j})}, [v]_{j_1}) : (u,v) \in \bar{\Box}(\vec{j})\}$. If

$\bar{\Box}(\vec{j}) \subseteq W \times 2^W \times \ldots \times 2^W$ then $\bar{\bar{\Box}}(\vec{j}) = \{([u]_{r(\vec{j})}, S_1/j_1, \ldots, S_n/j_n) : (u, S_1, \ldots, S_n) \in \bar{\Box}(\vec{j}) \}$. Let $\bar{C}' = \{\bar{\bar{\Box}}(\vec{j}) : \bar{\Box}(\vec{j}) \in \bar{C}\}$. (The restrictions on $\bar{\Box}(\vec{j})$ stemming from M's being a D-model insure that the members of $\bar{C}'$ are well defined.) Finally, let $V'(p^j) = \{[u]_j : u \in V(p^j)\}$ for all sort-j sentence letters p. It is now easy to check that $M' = (W', o', \bar{C}', V')$ is a C-model and that for all $[w]_j$ in W_j and all sort-j sentences, A^j, $(M', [w]_j) \models A^j$ iff $(M,w) \models A^j$. Therefore, if $\mathcal{M}' = \{M' : M \in \mathcal{M}\}$, $\models_{\mathcal{M}} = \models_{\mathcal{M}'}$. Furthermore, suppose $M = (W, o, \bar{C}, V) \in \mathcal{M}'$ and $w \in W_k$. Then there is some model $M^- = (W^-, E^-, o^-, \bar{C}^-, V^-)$ in $\mathcal{M}$ such that $(M^-)' = M$, and for some v in W^-, $w = \{u \in W^- : u \; E_j \; v\}$. Since $\mathcal{M}$ is normal, $(M^-)^v = (W^-, E^-, v, \bar{C}^-, V^-)$ is in $\mathcal{M}$. By our construction, then $((M^-)^v)' = (W, \{[v]_j : j \in Q\}, \bar{C}, V)$ is in $\mathcal{M}'$. Since $w = [v]_k$, w is in the designated set of $((M^-)^v)'$. Hence $\mathcal{M}'$ is semi-normal.

Lemmas 2.2 and 2.3 show that nothing substantive hinges on the choice between C-models and D-models. The D-models, however, highlight an important property of many-sorted frameworks. A D-model is just an ordinary modal model with special restrictions on the interpretations of connectives and the valuation function. Because the valuation function is restricted we cannot expect many-sorted logics to be closed under substitutions (compare Theorem 1.4f). Since a sentence letter of sort-j can be given a valuation which makes it true in exactly the same worlds as any j-sentence, however, we can expect many-sorted logics to be closed under substitution of sentences for sentence letters <u>of the same sort</u>. Partial failure of substitutivity, like difficulties in interpreting iterated modalities, is a sign that a many-sorted treatment might be appropriate.

Before proceeding to more technical results, a word should be said about what might have struck the reader as a needless complication in our account. According to the formulation given above an n-ary connective $\Box$ can apply to sentences of different sorts. Wouldn't it have been simpler to have each connective apply to only one kind of string and be interpreted by a single relation? The answer to this question is "Perhaps so, but it still wouldn't be desirable." Only "perhaps so," because if a connective had a type with an infinite domain it would have to be replaced by an infinite number of new connectives. "It still would not be desirable," because we would like the application of a connective to a sentence in our formal languages to mirror the application of connecting phrases to sentences in ordinary language. There are many cases in which a certain change in meaning in one kind of sentence parallels a change of meaning in another kind so closely that the same phrase is used to indicate both. For now we shall be content with a single example. Suppose S_1, S_2, S_3, S_4 are English sentences whose truth depends on features 1, 2, 3, 4, respectively. Then the sentence obtained by inserting 'and' between S_1 and S_2 depends on features 1 and 2. The sentence obtained by performing the same operation on S_2 and S_3 depends on features 3 and 4. Clearly the most natural many-sorted language which captures the procedure of inserting 'and' between sentences would contain a single connective, say '$\wedge$', that applies to sentences of sorts 1 and 2 (yielding a sentence of a new sort, say $\{1,2\}$) as well as to sentences of sorts 3 and 4 (yielding a sentence of a fourth sort, say $\{3,4\}$).

We shall hold to the idea, therefore, that a single connective of L may apply to many different kinds of strings of sentences. We shall occasionally, however, abandon this idea in order to bring out similarities between many-sorted and one-sorted frameworks. If M is a model suitable for L we call M' and L' the <u>simple versions</u> of M and L if L' results from L by replacing each type-r connective $\Box$ by connectives $\Box(\vec{j})$ of sort $r\restriction\{\vec{j}\}$ for each $\vec{j}$ in the domain of r, and M' is the model suitable for L' such that $W_{M'} = W_M$, $E_{M'} = E_M$, $o_{M'} = o_M$ and $\overline{\boxed{\vec{j}}}_{M'} = \overline{\Box}_M(\vec{j})$. L' and M' are "simple" in the sense that each connective of L' applies to a single kind of string and is interpreted in M' by a single relation. The usefulness of simple models in brought out by the following observation. If we ignore the sorts of the sentence letters and the types of the connectives we change a simple many-sorted language L' into a one-sorted language L^2 which contains as a subset all the sentences of the original language. If we drop the E-coordinate from a simple many-sorted model M' we obtain an ordinary model M^2 suitable for L^2. Furthermore, for all A in L', $M' \models A$ iff $M^2 \models A$. We call M^2 and L^2 the <u>unsorted versions</u> of M and L.

E. Completeness

Many-sorted languages and models form a wider class than do those discussed in Chapter I, but most of the notions defined earlier can be carried over to the many-sorted case with little or no change. For example, we can use the old definitions of 'consequence relation on L', 'regular consequence relation on L', and 'finitary consequence relation on L' exactly as they stand. The same applies to the definitions of '$\vdash$ consistent theory of L' and 'maximal $\vdash$ consistent theory of L', and to the lemmas concerning these notions (viz., Lemmas 1.9, 1.10, 1.11a). In the remainder of this chapter, terminology, notation, and results of the previous chapter which can be applied directly to the many-sorted case will occasionally be used without special comment. As before we take $\vdash$ to be a consequence relation on L.

In this section we show how the completeness proof in Chapter I can be altered to fit the many-sorted case. We prove two theorems (2.1 and 2.2) which are the counterparts of Theorems 1.4 and 1.5. First, however, we turn to the task of suitably generalizing the notion of classical model.

The salient features of a classical model are the simple relations between the truth values of sentences in the model and the truth values of their Boolean combinations. It is easy to show (in either the one-sorted or the many-sorted case) that these properties characterize the models associated with classical consequence relations. More precisely, let the definition of, say,

'$\vdash$ has classical conjunction' be altered to read "$\Gamma, A, B \vdash \Delta$ iff $\Gamma, A \wedge B \vdash \Delta$ <u>whenever</u> $A \wedge B$ <u>is a sentence</u>." Then it can be shown that if $\models_{\mathcal{M}}$ has classical conjunction, $M \in \mathcal{M}$ and $A \wedge B$ is a sentence, then $M \models A \wedge B$ iff $M \models A$ and $M \models B$. Conversely, if for every M in $\mathcal{M}$, $M \models A \wedge B$ iff $M \models A$, $M \models B$, and $A \wedge B$ is sentence then $\models_{\mathcal{M}}$ has classical conjunction. In the single-sorted case, however, we were able to do better than this. We were able to define a simple, fixed relation, Conj W, such that, whenever $\models_{\mathcal{M}}$ had classical conjunction, we could always assume '$\wedge$' was interpreted by Conj W in each member of $\mathcal{M}$ with domain W. Conj W, however, does not in general meet the special restrictions needed to be included in a many-sorted model. In order to fix the interpretation for the Boolean connectives in many-sorted models we need to pay a little more attention to sorts than we have done in this paragraph. The treatment below is motivated by the remarks at the end of Section E.

<u>Definition 1.5</u>. L is <u>partially f-Boolean</u>[2] if

(1) The members of Q are sets and $s \cup t \in Q$ whenever $s \in Q$ and $t \in Q$.

(2) If any of the symbols '$\wedge$', '$\vee$', '$\rightarrow$' are connectives of L, then they are of type r where, for all s and t in Q, $r(s,t) = s \cup t$.

---

[2] The 'f' stands for 'feature-dependent'.

(3) If the symbol '$-$' is a connective of L, then it is of type q where, for all s in Q, $q(s) = s$.

L is <u>f-Boolean</u> if it is partially f-Boolean and contains '$\wedge$', '$\vee$', '$\rightarrow$', and '$-$' as connectives.

<u>Definition 2.6</u>. $\vdash$ is <u>f-classical</u> (partially f-classical) if L is f-Boolean (partially f-Boolean) and (CC), (CD), (CI), (CD) of Definition 1.17 all hold.

<u>Definition 2.7</u>. Let $M = (W, E, o, \overline{C}, \overline{V})$ be a model suitable for the f-Boolean (partially f-Boolean) language L. Then M is f-classical (partially f-classical) if, for all s and t in Q, $E_{s \cup t} = E_s \cap E_t$ and each of the following conditions hold.

(a) $\overline{\wedge}(s,t) = \text{Conj } W(s,t) = \{(w,U,V) : [w]_{s \cup t} \in U/s \cup t \cap V/s \cup t\}$.

(b) $\overline{\vee}(s,t) = \text{Disj } W(s,t) = \{(w,U,V) : [w]_{s \cup t} \in U/s \cup t \cup V/s \cup t\}$.

(c) $\overline{\rightarrow}(s,t) = \text{Impl } W(s,t) = \{(w,U,V) : U/s \cup t \subseteq V/s \cup t\}$.

(d) $\overline{-}(s) = \text{Neg } W(s) = \{(w,U) : [w]_s \notin U/s \}$.

<u>Lemma 2.4</u>. If M is a partially f-classical model suitable for L then the following hold for all A, B in L and all points w of M.

(a) If $A \wedge B$ is a sentence of L, then $(M,w) \models A \wedge B$ iff $(M,w) \models A$ and $(M,w) \models B$.

(b) If $A \vee B$ is a sentence of L, then $(M,w) = A \vee B$ iff $(M,w) \models A$ or $(M,w) \models B$.

(c) If $A \to B$ is a sentence of L, then $(M,w) \models A \to B$ iff $(M,w) \not\models A$ or $(M,w) \models B$.

(d) If $-A$ is a sentence of L, then $(M,w) \models -A$ iff $(M,w) \not\models A$.

The proof is easy. We do (a) as an example. Suppose $s(A) = s$, $s(B) = t$. Then $(M,w) \models A \wedge B$ iff Conj $W_M(s,t)$ w, $|A|_M$, $|B|_M$, i.e., iff $[w]_{s \cup t} \in |A|_M/s \cup t \cap |B|_M/s \cup t$. Therefore $(M,w) \models A \wedge B$ implies $[w]_{s \cup t} \in |A|_M/s \cup t$ which implies there is a v in W such that $w\ E_s\ v$ and $w\ E_t\ v$ and $v \in |A|_M$. Since A is sort s, Lemma 2.1 says that $(M,w) \models A$. Similarly, $(M,w) \models A \wedge B$ implies $(M,w) \models B$. Now suppose $(M,w) \models A$ and $(M,w) \models B$; i.e., $w \in |A|_M \cap |B|_M$. Then clearly $[w]_q \in |A|_M/q \cap |B|_M/q$ for all sorts q. Hence $(M,w) \models A\ \ B$.

The classification of the non-Boolean connectives with respect to $\vdash$ is straightforward.

<u>Definition 2.8</u>. Suppose $\vdash$ is a consequence relation on L, $\Box$ and $\boxminus$ are type-r connectives of L, and $\vec{j} = (j_1,\ldots,j_n)$ is in the domain of r.

a) $\Box$ is <u>$\vec{j}$-Montague with respect to $\vdash$</u> if, for all $A_1,\ldots,A_n$ and all $B_1,\ldots,B_n$ in L, if $s(A_i) = s(B_i) = j_i$ for $1 \leq i \leq n$ then M of page holds.

b) $\Box$ is <u>$\vec{j}$-normal-Kripke</u> (<u>$\vec{j}$-normal-K4</u>; <u>$\vec{j}$-normal-S4</u>; <u>$\vec{j}$-normal-S5</u>) <u>with respect to $\vdash$</u> if, for all subsets Γ, Δ, Ψ, Θ of L and all A, B in L, if $s(\Gamma) = s(\Delta) = s(A) = s(B) = j$

and $\vec{j} = (j)$ then $K_{\Box}$ ($K_{\Box}$ and $4_{\Box}$; $K_{\Box}$, $4_{\Box}$, $S_{\Box}$; $K_{\Box}$, $4_{\Box}$, $S_{\Box}$, $5_{\Box}$) on page hold.

c) $(\Box, \boxminus)$ is a pair of $\vec{j}$-<u>tense connectives with respect to</u> $\vdash$ if, for all Γ, Δ, B meeting the conditions above, $T_{\Box\boxminus}$ holds.

$\Box$ is <u>Montague</u> (<u>normal-Kripke</u>; <u>normal-K4</u>; <u>normal-S4</u>; <u>normal-S5</u>) with respect to $\vdash$ if all instances of $M_{\Box}$ ($K_{\Box}$ and $4_{\Box}$; $K_{\Box}$, $4_{\Box}$, $S_{\Box}$; $K_{\Box}$, $4_{\Box}$, $S_{\Box}$, $5_{\Box}$) which are well-formed hold. (So $\Box$'s being Montague entails its being $\vec{j}$-Montague for all $\vec{j}$ in the domain of r, but the converse of this may fail.)

<u>Theorem 2.1</u>. Let M be a normal class of models suitable for L. Then

a) Every connective of L is $\vec{j}$-Montague with respect to $\models_M$.

b) If every member of M is partially f-classical then $\models_M$ is partially f-classical.

c) If, for all $(W, E, o, \ldots, \bar{\Box}(\vec{j}), \ldots, V)$ in M, $\bar{\Box}(\vec{j})$ is a point-to-point (transitive point-to-point; transitive and reflexive point-to-point; transitive reflexive and symmetric point-to-point) relation then $\Box$ is a $\vec{j}$-normal-Kripke, $\vec{j}$-normal-K4; $\vec{j}$-normal-S4; $\vec{j}$-normal-S5 connective with respect to $\models_M$.

d) If, for all $(W, E, o, \ldots, \bar{\Box}(\vec{j}), \bar{\boxminus}(\vec{j}), \ldots, V)$ in M, $\bar{\Box}(\vec{j})$ and $\bar{\boxminus}(\vec{j})$ are transitive, reflexive point-to-point relations such that $\bar{\square}(\vec{j})$ is the converse of $\bar{\boxminus}(\vec{j})$ then $(\square, \boxminus)$ is a pair of tense connectives with respect to $\models_{M}$.

Proof. The proof is almost identical to that of Theorem 1.4. We do subcase (i) of (a) as an example. Let $\models\ =\ \models_M$. Suppose $\square$ is type r, $A_i \dashv\models B_i$ and $s(A_i) = s(B_i) = j_i$ for $1 \leq i \leq n$, $r(j_1,\ldots,j_n) = j_o$, and finally $M = (W, E, o,\ldots,\bar{\square}(j_1,\ldots,j_n),\ldots,V)$ is a model in M such that $R = \bar{\square}(j_1,\ldots,j_n) \subseteq W \times 2^W \times \ldots \times 2^W$. Then $M \models \square A_1\ldots A_n$ iff $R\ o, |A_1|_M,\ldots,|A_n|_M$. Since M is normal and $A_i \dashv\models B_i$ for $1 \leq i \leq n$, each $|A_i|_M$ equals $|B_i|_M$. Hence $M \models \square A_1\ldots A_n$ iff $M \models \square B_1\ldots B_n$. Therefore $\square A_1\ldots A_n \dashv\models \square B_1\ldots B_n$.

Theorem 2.2. If $\vdash$ is a Montague consequence relation on L (see Definition 1.9), then there is a normal class M of models suitable for L such that $\vdash$ is strongly complete for M and the following hold:

a) If $\vdash$ is partially f-classical then each member of M is partially f-classical.

b) If $\square$ is $\vec{j}$-normal-Kripke ($\vec{j}$-normal-K4, $\vec{j}$-normal-S4) with respect to $\vdash$ then, for all $(W, E, o,\ldots, \bar{\square}(\vec{j}),\ldots, V)$ in M, $\bar{\square}(\vec{j})$ is a point-to-point (transitive point-to-point, transitive and reflexive point-to-point) relation.

c) If $\vdash$ has classical negation and $\square$ is $\vec{j}$-normal-S5 with respect to $\vdash$, then, for all $(W, E, o,\ldots, \bar{\square}(\vec{j}),\ldots, V)$ in

M, $\bar{\Box}(\vec{j})$ is a transitive, reflexive and symmetric point-to-point relation.

d) If $\vdash$ has classical negation, then if $(\Box, \boxminus)$ is a pair of $\vec{j}$-tense connectives with respect to $\vdash$ then, for all $(W, E, o, \ldots, \bar{\Box}(\vec{j}), \bar{\boxminus}(\vec{j}), \ldots, V)$ in M, $\bar{\Box}(\vec{j})$ and $\bar{\boxminus}(\vec{j})$ are transitive and reflexive and $\bar{\Box}(\vec{j})$ is the converse of $\bar{\boxminus}(\vec{j})$.

To prove this we shall need a few definitions and lemmas.

Lemma 2.5. Suppose the following all hold.

(1) $\Box$ is a type-r connective of L and $(j_1, \ldots, j_n)$ is in the domain of r.

(2) M and M' are two models suitable for L which are identical except that M has $\bar{\Box}(j_1, \ldots, j_n)$ where M' has $\bar{\Box}'(j_1, \ldots, j_n)$.

(3) For all subsets $U_1, \ldots, U_n$ of W_M, if $U_1 = |A_1|_M, \ldots, U_n = |A_n|_M$ for some $A_1, \ldots, A_n$ in $L^{j_1}, \ldots, L^{j_n}$ respectively and if $w \in L^{r(j_1, \ldots, j_n)}$ then $\bar{\Box}\, w, U_1, \ldots, U_n$ iff $\bar{\Box}'\, w, U_1, \ldots, U_n$.

Then for all sentences A of L, $M \models A$ iff $M' \models A$.

Proof. Straightforward induction on the length of A.

Definition 2.9. A canonical model for $\vdash$ is a 5-tuple $(W, E, o, \bar{C}, V)$ such that:

(1) $W = \|W\|_{\vdash}$ is the set of all maximal $\vdash$ consistent sets of

sentences of L.

(2) $E = \{\|E_j\|_\vdash : j$ is a sort of $L\}$ where $\|E_j\|_\vdash =$ $\{(u,v) \in \|W\|_\vdash \times \|W\|_\vdash : u \cap L^j = v \cap L^j\}$.

(3) $o \in \|W\|_\vdash$.

(4) $\overline{C} = \{\| \Box(j_1,\ldots,j_n)\|_\vdash : \Box$ is a type r connective of L and $(j_1,\ldots,j_n)$ is in the domain of $r\}$, where $\| \Box(j_1,\ldots,j_n\|_\vdash$ is defined as follows: If $\Box$ is j_1-Kripke with respect to $\vdash$ then $\| \Box(j_1)\|_\vdash = \{(u,v)$: for all A in L^{j_1}, $\Box A \in u \Rightarrow A \in v\}$. If $\Box$ is not j_1-Kripke with respect to $\vdash$, then $\| \Box(j_1,\ldots,j_n)\|_\vdash = \{(w, U_1,\ldots,U_n)$: for some $A_1,\ldots,A_n$ of sort $j_1,\ldots,j_n$ respectively, $U_1/j_1 = \|A_1\|_\vdash/j_1,\ldots, U_n/j_n = \|A_n\|_\vdash/j_n$ and $A_1\ldots A_n \in w\}$.

(5) V is a function from the sentence letters of L into the subsets of W such that $V(p^j) = \{w : p^j \in w\}$.

Lemma 2.6. If $M = (W, E, o, \overline{C}, V)$ is a canonical model for $\vdash$, then it is a model suitable for L.

Proof. We need only verify that the proper restrictions on V and the members of $\overline{C}$ all hold. First, if $w\ E_j\ w'$ then $w \cap L^j = w' \cap L^j$. Hence $p^j \in w$ iff $p^j \in w'$, i.e., $w \in V(p^j)$ iff $w' \in V(p^j)$. Second, suppose $\Box$ is a type-r Kripke connective, j is in the domain of r, $u\ E_{r(j)}\ u'$ and $v\ E_j\ v'$. Then $(u,v) \in \| \Box(j)\|_\vdash$ iff for all A in L^j, $A \in u$ implies $A \in v$. But $u\ E_{r(j)}\ u'$ means $u \cap L^{r(j)} = u' \cap L^{r(j)}$ and A in L^j

implies $\Box A$ in $L^{r(j)}$. Hence $\Box A \in u$ iff $\Box A \in u'$. Similarly $A \in v$ iff $A \in v'$. Therefore $(u,v) \in \|\Box(j)\|_{\vdash}$ iff for all A in L^j, $\Box A \in u'$ implies $A \in v'$, i.e., iff $(u',v') \in \|\Box(j)\|_{\vdash}$. Finally, suppose $\Box$ is a type-r Montague connective which is not Kripke, $(j_1,\ldots,j_n)$ is in the domain of r, $u\ E_j\ v$ where $j = r(j_1,\ldots,j_n)$ and, for $1 \leq i \leq n$, $U_1/j_i = V_i/j_i$. Then $(u, U_1,\ldots,U_n)$ is in $\|\Box(j_1,\ldots,j_n)\|$ iff for some $A_1,\ldots,A_n$ in $L^{j1},\ldots,L^{jn}$ respectively, $U_1/j_1 = \|A_1\|_{\vdash}/j_1,\ldots, U_n/j_n = \|A_n\|_{\vdash}/j_n$ and $\Box A_1\ldots A_n \in u$. Since $u \cap L^j = v \cap L^j$, and $U_i/j_i = V_i/j_i$, this holds if $(v, V_1,\ldots,V_n)$ is in $\|\Box(j_1,\ldots,j_n)\|$.

<u>Lemma 2.7</u>. Suppose $s(A) = s(B) = j$. Then $\|A\|_{\vdash}/j \subseteq \|B\|_{\vdash}/j$ iff $A \vdash B$.

<u>Proof</u>. First suppose $A \vdash B$. Then $[u]_j \in \|A\|_{\vdash}/j \Rightarrow$ there is a v in $\|W\|_{\vdash}$ such that $u\ E_j\ v$ and $v \in \|A\|_{\vdash}$, i.e., there is a v in $\|W\|_{\vdash}$ such that $u \cap L^j = v \cap L^j$ and $A \in v$. Since $A \vdash B$ and v is maximal $\vdash$ consistent, $B \in v$. But B is of sort-j, so $B \in u$. Hence $u \in \|B\|_{\vdash}$ and $[u]_j \in \|B\|/j$. Next suppose $A \nvdash B$. Then $u = (\{A\}, \{B\})$ has a maximal consistent extension, u' and $[u']_j \in \|A\|_{\vdash}/j$ but $[u']_j \notin \|B\|/j$.

<u>Lemma 2.8</u>. If $M = (W, E, o, \bar{C}, V)$ is a canonical model for $\vdash$ and A is in L then $(M,w) = A^-$ iff $A \in w$.

<u>Proof</u>. By induction (similar to the proof of Lemma 1.12)

- If A is a sentence letter, $(M,w) = A$ iff $w \in V(A)$ iff $A \in w$.

- If $A = \Box B$ where $\Box$ is (j)-Kripke and B is sort-j, then $(M,w) \not\models A$ iff, for some u in W, $w \,\|\Box(j)\|_{\vdash}\, u$ and $(M,u) \not\models B$. By induction hypothesis, then, $B \notin u$. But B is sort-j, so this means $\Box B$ cannot be in W. Conversely, if $\Box B \notin w$ then the theory $u = (\{C \in L^j : \Box C \in w\}, \{B\})$ must be consistent. (If it weren't then the (j)-normal-Kripke-ness of $\vdash$ would be violated.) Therefore we can extend u to a maximal $\vdash$ consistent theory u' such that $w \,\|\Box(j)\|_{\vdash}\, u'$. But $B \notin u'$. Hence by the induction hypothesis this means $(M,w) \models \Box B$.

- If $A = \Box B_1 \ldots B_n$ where $\Box$ is $\vec{j}$-Montague and not $\vec{j}$-Kripke with respect to $\vdash$ and $\vec{j} = (s(B_1),\ldots,s(B_n))$ then $(M,w) \models A$ iff $\|\Box(\vec{j})\|_{\vdash}\, w, \|B_1\|_M,\ldots,\|B_n\|_M$. By induction hypothesis this holds iff $\|\bar{\Box}(j)\|_{\vdash}\, w, \|B_1\|_{\vdash},\ldots,\|B_n\|_{\vdash}$. Therefore if $(M,w) \neq A$, $\|\bar{\Box}(\vec{j})\|_{\vdash}\, w, \|B_1\|_{\vdash},\ldots,\|B_n\|_{\vdash}$ does not hold. Hence $\Box B_1 \ldots B_n$ is not in w. Conversely, if $(M,w) \models A$ then $\bar{\Box}\, w, \|B_1\|_{\vdash},\ldots,\|B_n\|_{\vdash}$, so there are sentences $C_1,\ldots,C_n$ such that for $1 \leq i \leq n$, $\|C_i\|_{\vdash}/j_i = \|B_i\|_{\vdash}/j_i$ and $C_1 \ldots C_n$ is in w. By Lemma 2.7, $C_1 \dashv\vdash B_i$ for $1 \leq i \leq n$. Since $\Box$ is τ-Montague, $\bar{\Box} C_1 \ldots C_n \dashv\vdash \bar{\Box} B_1 \ldots B_n$. Hence $A \models \Box B_1 \ldots B_n$ is in w.

Theorem 2.2 can now be proved from the above lemmas exactly as Theorem 1.5 was proved from the lemmas preceding it. More specifically, let M^- be the class of all canonical models for L,

M be the result of replacing $\overline{\wedge}(s,t)$, $\overline{\vee}(s,t)$ $\overline{\rightarrow}(s,t)$, $\overline{-}(s,t)$ in each model of M^- first coordinate W by Conj W(s,t), Disj W(s,t), Impl W(s,t), Neg W(s,t), respectively, then it is easy to check that:

(1) $\vdash$ is complete for M.

(2) If $\Box$ is $\vec{j}$-normal-Kripke ($\vec{j}$-normal-K4, $\vec{j}$-normal-S4) with respect to $\vdash$ then $\overline{\Box}(\vec{j})$ is a point-to-point (transitive point-to-point, transitive and reflexive point-to-point relation).

(3) If L contains the unary connective '-', CN of Definition 1.17 holds and $\Box$ is $\vec{j}$-normal-S5 with respect to $\vdash$, then $\overline{\Box}(\vec{j})$ is a transitive , reflexive and symmetric point-to-point relation.

F. Reduction to single sort

It is well-known that classical, many-sorted theories are reducible to single sorted ones. There is a trivial analog to this in the case of our many-sorted propositional languages. For there is a natural translation from a language L to its one-sorted version L^2, namely, the function f such that $f(p) = p$ for all sentence letters p and $f(\Box A_1 \ldots A_n) = \boxed{\vec{j}} A_1 \ldots A_n$ where $\vec{j} = (s(A_1),\ldots,s(A_n))$. Since f is one-one the f-extension $\vdash'$ of $\vdash$ must be a consequence relation on L' which has $\vdash$ as a fragment (see remarks after Definition 1.40).

We cannot assume, however, that properties of $\vdash$ carry over to $\vdash'$. For example, suppose $\Box$ is a unary connective of L which is Montague with respect to $\vdash$ and which applies to A, but not to B; suppose further that $A \dashv\vdash B$ and not $A \vdash \phi$. Since $\Box B \notin L$ we do not have $\boxed{\vec{j}}\, A \vdash' \boxed{\vec{j}}\, B$ where $\vec{j} = (s(A))$, so $\boxed{\vec{j}}$ is <u>not</u> Montague with respect to $\vdash'$.

We would like to modify the definition of $\vdash'$ so that $\boxed{\vec{j}}\, A \vdash' \boxed{\vec{j}}\, B$ holds, but we must be careful that in so doing we do not get $\boxed{\vec{j}}\, B \vdash' C$ for any C in L such that $\Box A \nvdash C$. And there are properties besides $\Box$'s Montagueness that we would like to preserve as well. In this section a construction suggested by the standard reduction of many-sorted predicate systems is used to obtain a more suitable $\vdash'$. We borrow freely from the first part of section H, Chapter I. (This is permissible because a language is taken there to be merely a collection of structureless objects.)

Let L^{o} be the language obtained by adding to the unsorted version of L a new unary connective ⓙ for each sort-j in Q. (Intuitively ⓙ asserts something like "A is a true sort-j sentence" and, if 'v' and '-' are classical, ⓙA $\vee$ ⓙ- A asserts something like "A is a sort-j sentence." But this reading is not quite accurate because ⓙA will turn out to be valid whenever A is.)

Let $f : L \to L^{o}$ be defined as follows: If A is a sort-j sentence letter, $f(A) =$ ⓙA; if $A = \Box A_1 \ldots A_n$ where $\Box$ is a type-r connective of L, $\vec{j} = (s(A_1), \ldots, s(A_n))$ and $r(\vec{j}) = j$, then $f(A) =$ ⓙ $\boxed{\vec{j}}\, f(A_1) \ldots f(A_n)$. In order to make certain that the consequence relation we construct has the right properties we must now make a detour through models.

If M is a model suitable for L, call a model M^o suitable for L^o a one-sorted derivative of M if $W_{M^o} = W_M$, $o_{M^o} = o_M$, $\boxed{T}_{M^o} = \bar{\Box}_M(\vec{j})$ for all $\Box$ in L, $\textcircled{T}_{M^o} = E_{j_M}$, and V_{M^o} is such that, if p is a sort-j sentence letter of L, then $u \in V_M(p)$ iff for all w, $u\, E_{j_M}\, w$ implies $w \in V_{M^o}(p)$. (Notice that each model suitable for L^o is a derivative of at most one model suitable for L. An easy induction establishes that if M^o is a one-sorted derivative of M, then for all A in L, $(M^o,w) \models f(A)$ iff $(M,w) \models A$.

By Theorem 2.2 we know $\vdash$ is complete for a class of models $\mathcal{M}$ such that, for all M in $\mathcal{M}$, $\bar{\Box}_M(\vec{j})$ is a point-to-point (transitive point-to-point, etc.) relation if $\Box$ is $\vec{j}$-Kripke ($\vec{j}$-K4 etc) and $\bar{\wedge}_M(\vec{j})$ $(\bar{-}(\vec{j})$, etc) is Conj $W(\vec{j})$ (Neg $W(\vec{j})$, etc) if $\vdash$ has $\vec{j}$-classical conjunction ($\vec{j}$-classical negation, etc). Let $\mathcal{M}^o$ be the class of all M^o such that, for some M in $\mathcal{M}$, M^o is a derivative of M. From the observation in the preceding paragraph it follows that, for all $\Gamma,\Delta \subseteq L$, $\Gamma \models_{\mathcal{M}} \Delta$ iff $f(\Gamma) \models_{\mathcal{M}^o} f(\Delta)$, so $\models_{\mathcal{M}}$ is a fragment of $\models_{\mathcal{M}^o}$.

It remains to show that $\models_{\mathcal{M}^o}$ is more useful than the $\vdash'$ defined earlier. Notice that the relevant properties of $\Box$ with respect to $\vdash$ are characterized by conditions on the relations $\bar{\Box}_M(\vec{j})$. But, for each $\vec{j}$ and each M in $\mathcal{M}$, $\bar{\Box}_M(\vec{j}) = \boxed{T}_{M^o}$. Hence we can infer immediately that if $\Box$ is $\vec{j}$-Kripke with respect to $\vdash$ ($= \models_{M}$), for example, then $\boxed{j}$ is Kripke with respect to $\models_{M^o}$. (The special connectives ⓙ are normal-S5 with respect to $\models_{M^o}$ because $\bar{ⓙ}$ is always an equivalence relation.) It will always be clear from the context what kind of models the arguments and values of g_1^* and g_2^* are intended to be. We can show that is valuation-unrestricted. (in the sense that it contains all models $(W, E, o, \bar{C}, V)$ suitable for L such that $(W, E, o, \bar{C}, U)$ is in M for some U) if and only if M^o is valuation-unrestricted (in the usual sense). For suppose M is valuation-unrestricted $M^o = (W, o, \bar{C}, V)$ is in M^o and X is a function from sentence letters to subsets of W. Then there is some model $M = (W, E, o, \bar{B}, U)$ such that M^o is a derivative of M. If p is a sort-j sentence letter of L, let $U'(p) = \{w : \forall u\ (u\ E_j\ w \Rightarrow u \in X(p))\}$. Then (W, E, o, B, U') is in M. But $(W, o, \bar{C}, X)$ is a derivative of $(W, E, o, \bar{B}, U')$. Hence it is in M^o. To prove the converse it is sufficient to note that every model M in M has a derivative M^o in M^o with the same valuation. Furthermore, if M^o is also a derivative of M', then M = M'.

We now collect for reference some facts which follows from the above.

Theorem 2.3. Let $\vdash$ is a consequence relation on L and L^o is the (one-sorted) language obtained by adding to the unsorted version of L a new unary connective $\boxed{j}$ for each j in Q.

Then there is a consequence relation $\vdash^{o}$ on L^{o} such that:

1) $\vdash$ is a fragment of $\vdash^{o}$.

2) If $\vdash$ is normal-Kripke (normal-K4, normal-S4, normal-S5) then so is $\vdash^{o}$.

3) If $\vdash$ is complete for a class of frames then so is $\vdash^{o}$.

G. Sort restrictions in the standard logics

At the end of section G of Chapter I it was shown that completeness results for many simple modal logics follow from Theorems 1.4 and 1.5. Similarly, we can now use Theorems 2.1 and 2.2 to obtain completeness results for some many-sorted versions of these logics. But first we should say something about why sort distinctions are needed. The most convincing arguments for sorts, it seems to me, occur in connection with tense logics. These are discussed in section C. A case can be made, however, for the appropriateness of sort distinctions in logics for necessity.

Sentences, it might be argued, come in two varieties. Those of the first variety, (e.g., "The cat is on the mat") are not true or false but true or false _in a particular situation_. Now suppose we prefix to a sentence of the first variety the phrase "There is some situation in which...". The resulting sentence is of a new variety -- it is not true-in-a-situation or false-in-a-situation, but simply 'true' or 'false'. But it follows from our earlier remarks that the operation of prefixing "There is some situation in which..." to an English sentence is mirrored in certain of our formal frameworks by the operation of prefixing the connective $\Box$ to a sentence. These frameworks are ones in which the language contains $\Box$ as the only non-Boolean connective and $\Box$ is normal-S5. We can therefore distinguish between at least two sorts of sentences, those which are situation-evaluated (sort $\{1\}$) and those which are not (sort $\{2\}$).[3] $\Box$ applies to sort $\{1\}$ sentences and yields sort $\{2\}$ sentences. Since it does not make sense to prefix "There is a situation in which..." to sentences which are not situation-evaluated, we can also specify that $\Box$ does not apply to sort-$\{2\}$ sentences at all. We can introduce a new sort, $\{1,2\}$, to include Boolean combinations of sentences which are situation-evaluated and sentences which are not. It seems reasonable to allow $\Box$ to apply to sort-$\{1,2\}$ sentences as well. The result will always be a sort-$\{2\}$ sentence. Call the language with these features L^*. Let $\vdash^*_{S5}$ be the smallest classical consequence relation -- on L^* such that $\Box$ is normal-S5 with respect to $\vdash$. (We could similarly define $\vdash^*_K$, $\vdash^*_{K4}$, $\vdash^*_{S4}$, but there seems little reason to do so.) Let M^* be the class of all classical models suitable for L^* such that, for all M in M^*, $(E_{\{1,2\}})_M = (E_{\{1\}})_M = \{(u,v):u = v\}$, $(E_{\{2\}})_M = W_M \times W_M$, and $\bar{\Box}_M((\{1\})) = \bar{\Box}_M((\{1,2\})) = W_M \times W_M$.

---

[3] The fact that a sentence is situation-evaluated does not mean its truth must change from one situation to another. We want to distinguish between sentences like '$p \vee - p$' and '$p \wedge - p$' which happen to come out true in every situation or false in every situation and sentences like '$\Box p$' in which the present situation plays no role in the assessment of truth.

Theorem 2.4. $\vdash^{*}_{S5} = \models_{M^*}$.

Proof.

(i) Soundness. It is sufficient to prove that $\models_{M^*}$ is classical and closed under all well-formed instances of $K_{\Box}$, $4_{\Box}$, $5_{\Box}$, and $S_{\Box}$. For, since $\vdash^{*}_{S5}$ is the smallest such consequence relation this will show $\vdash^{*}_{S5} \subseteq \models_{M^*}$. These facts can be proved in the same manner as in the one-sorted case. (Theorem 1.4 c,d) For example we prove $K_{\Box}$: Suppose $\Gamma \models_{M^*} A$ where $\Gamma \cup \{A\} \subseteq L^{\{1\}} \cup L^{\{1,2\}}$ and $M \in M^*$ such that $M \models \Box\Gamma$. Then, for all B in $L^{\{1\}} \cap \Gamma$ and all u such that $o \bar{\Box}_M((\{1\}))\ u$, $(M,u) = B$. Similarly, for all B in $L^{\{1,2\}} \cap \Gamma$ and all u such that $o \bar{\Box}_M((\{1,2\}))\ u$, $(M,u) \models B$. These conditions together just say that for all u in W_M and all B in Γ $(M,u) \models B$. Since $\Gamma \models_{M^*} A$ this means $M \models \Box A$ (regardless of whether A is in $L^{\{1\}}$ or $L^{\{1,2\}}$).

(ii) Sufficiency. Suppose not $\Gamma \vdash^{*}_{S5} \Delta$. Then there is a canonical model M for $\vdash^{*}_{S5}$ such that $M \models \Gamma$ but, for all B in Δ, $M \not\models B$. Let M^* be the member of M^* with domain $W^* = \{u : o\ \|\Box\|_{\vdash}\ u\}$ and valuation V^* such that for all sentence letters p, $V^*(p) = V_M(p) \cap W^*$.

Claim: For all w in W_{M^*}, $(M^*,w) \models A$ iff $(M,w) \models A$.

Proof. By induction on A. The crucial case is $A = \Box B$ for $B \in L^{\{1,2\}}$, $(M,w) \models A$ iff $\Box B \in w$ (by Lemma 2.8). Since

$B \underset{S5}{\overset{*}{\dashv}}\underset{S5}{\overset{*}{\vdash}} B \wedge (\sim \Box B \vee \Box B)$, we have $\Box B \underset{S5}{\overset{*}{\dashv}}\underset{S5}{\overset{*}{\vdash}} \Box(B \wedge (\sim \Box B \vee \Box B))$.

So $\Box B \in w$ iff $\Box(B \wedge (\sim \Box B \vee \Box B)) \in w$ iff $(M,w) \models \Box(B \wedge (\sim \Box B \vee \Box B))$ iff $\forall v \in W_M$ $w \parallel ((\{1,2\}))\parallel v$ implies $(M,v) \models B \wedge (\sim \Box B \vee \Box B)$ iff $\forall v \in W_{M^*}$, $(M,v) \models B$ and $(M,v) \models (\sim \Box B \vee \Box B)$ iff $\forall v \in W_{M^*}$, $(M,v) \models B$ iff (by induction hypothesis) $\forall v \in W_{M^*}$ $(M^*,v) \models B$ iff $(M^*,v) \models \Box B$. This proves the claim. Hence $M^* \models \Gamma$ and, for all B in Δ, $M^* \not\models B$. Thus $\Gamma \not\models_{M^*} \Delta$ and the theorem is proved.

<u>Corollary</u>. $\mathcal{L}(\vdash^*_{S5}) = S5 \cap L^*$.

Proof. Let $\mathcal{M}$ be the class of all (one-sorted) models, M, suitable for the (one-sorted) language L with the sentence letters of L^*, binary connectives '∧', '∨', '→', and unary connectives '-' and '□' such that M is classical and $\bar{\Box}_M = W_M \times W_M$. From Chapter I we know $S5 = \mathcal{L}(\models_{\mathcal{M}})$. For all M in $\mathcal{M}$ let $M^+ = (W, E, o, \bar{\Box}((\{1\})), \bar{\Box}((\{1,2\})), V)$ be the classical model suitable for L^* such that $W = W_M$, $E_{\{1\}} = E_{\{1,2\}}$ = equality on W, $E_{\{2\}} = W \times W$, $o = o_M$, $V = V_M$ and $\bar{\Box}((\{1\})) = \bar{\Box}((\{1,2\})) = W \times W$. $\mathcal{M}^* = \{M^+ : M \in \mathcal{M}\}$ is just $\mathcal{M}^*$. Furthermore, for all A in L^* and all M in $\mathcal{M}$, $(M,w) \models A$ iff $(M^+,w) \models A$. Hence $\vdash^*_{S5} = \models_{\mathcal{M}} \cap (2^{L^*} \times 2^{L^*})$ and $\mathcal{L}(\vdash^*_{S5}) = S5 \cap L^*$.

<u>Corollary</u>. $\vdash^*_{S5}$ is complete for a class of frames.

CHAPTER III

SORTS AND THE TENSES OF ENGLISH

A. Introduction

It has often been observed that "tense connectives" like those discussed in Chapter I do not provide a very satisfactory analysis of the tense constructions of English. In this section we try to show that all the common English tense constructions can be represented in a many-sorted tense logic. More specifically, we construct a system called ET (for 'English Tenses') in which sort distinctions are used to repair at least three flaws in the representation of English sentences by formulas of traditional tense logics.

Perhaps the most obvious discrepancy between tensed English sentences and formulas[1] of a one-sorted tense logic is that the former may describe events which take place over a period of time, whereas the latter are always evaluated at points representing instants of time. When we say "John built a house," for example, we are talking about what happened at an interval that includes both the moments John spent laying the foundation and those he spent

---

[1] In this chapter we depart from the terminology of the rest of the thesis by using the word 'formula' only in connection with formal languages and the word 'sentence' only in connection with natural languages.

shingling the roof. (To describe the state of affairs at a particular instant of this interval we use the special construction 'John was building a house'.) On the other hand, sentences like 'DeGama crossed the equator' and 'The cat was on the mat' describe states of affairs which obtained at a single moment of time.

A second deficiency of one-sorted tense logics is their failure to take account of the differences between sentences like 'John swims in the Channel' and 'John swims across the Channel'. The first sentence has the property that it truly describes an event which takes place during an interval I of time if and only if it truly describes an event which takes place at all subintervals of I. The second does not. A distinction like this is needed to account for the fact that 'John is swimming in the Channel' implies 'John has swum in the Channel', but 'John is swimming across the Channel' does not imply 'John has swum across the Channel'. This sort of phenomenon has been noted by grammarians and philosophers since Aristotle.[2]

A third shortcoming of the traditional tense logics is that formulas are evaluated only relative to their <u>utterance-times</u>. In natural languages the context in which a sentence is uttered often indicates that it should be evaluated relative to some other time. The formula representing 'Baltimore won the Pennant', for example, is true at time t if there is some time preceding t at which 'Baltimore wins the Pennant' is ture. But a sentence like

---

[2] See [Penner 1970].

'Baltimore won the Pennant' is usually uttered with a specific time in mind and is considered false unless the event described took place at that time. In other words, past tense sentences are generally taken to be making an implicit reference to a specific prior time. To express the idea that there was some prior time at which the event occurred it is more natural to use the present perfect, e.g., 'Baltimore has won the Pennant'). In traditional tense logics, however, there is no way to distinguish between simple past and present perfect.

The first two of these problems can be solved without much difficulty. To deal with the first, we require that the language of ET contain two sorts of sentences: 'interval-evaluated' and 'instant-evaluated' ones. The progressive tense operator will apply only to interval-evaluated sentences, and yield only instant-evaluated ones. To deal with the second problem, we distinguish between two kinds of interval-evaluated sentences: those with the subinterval property and those without it and specify that, if p is an interval-evaluated sentence letter with the subinterval property, then any ET-model which makes p true at interval I must make p true at all subintervals of I.

The third problem, however, is not so easy to handle. We can't simply ignore the utterance time and evaluate 'Baltimore won the Pennant' relative to an arbitrary reference time. For we do not want to accept the sentence as true unless the time referred to precedes the utterance time. Uttered today in a discussion about the events of 1980, the sentence is simply inappropriate. For this reason it seems natural to evaluate sentences relative to <u>two</u> times -- an utterance time and a reference time . Then 'Baltimore won the Pennant' will be true when uttered at time t with reference to t' if t' precedes t and 'Baltimore wins the Pennant' is true at t'. Several authors[3] have in fact argued that this kind of 'two-dimensionality' is needed to deal with tenses in subordinate clauses anyway, so this step is not so drastic. Unfortunately, however, it is not enough. For if the Boolean connectives are interpreted classically, then it is apparent that Boolean combinations of future tense and past tense sentences will always turn out to be inappropriate. These are circumstances, however, in which a sentence like 'If Robinson missed batting practice, Baltimore will lose' seems perfectly appropriate.

Apparently compound sentences can make implicit reference to <u>several</u> times. If we are to incorporate this feature in ET, we must specify some procedure for determining which time is relevant to the truth of which clause. One solution would be to evaluate formulas at <u>sequences</u> of times. A conditional $p \supset q$, for example, would be true at (u,v) if q were true at v whenever p was true at u. There are, however, two objections to this approach. First, it requires us to consider arbitrarily long sequences of times. Second, it allows us to assign distinct truth values to classically equivalent formulas. $p \wedge q$ and $-r \supset -s$, for example, could both be false at (u,v) even though $q \wedge p$ and

---

[3]See [Kamp 1971], [Vlatch 1973], and [Gabbay 1974].

$s \supset r$ were both true.

It seems preferable, therefore, to say that we actually determine which time a clause refers to by noticing its tense. We assume the speakers remarks are appropriate and true and we try to choose from among several possible alternatives a reference time which will verify this assumption. In ET, therefore, formulas are evaluated at utterance times with respect to sets of possible reference times. A past tense formula $P\,q$ is true at t with reference to R if there is some (instant or interval) t' in R such that t' precedes t and q is true at t'. $A \wedge B$ is true at t with reference to a set R if both A and B are true at t with reference to R. For example, suppose the context indicates that the only possible reference times are the intervals and instants of June 1, 1970. The sentence 'Robinson took batting practice and Baltimore will lose', uttered at gametime on June 1, is considered true if there is some interval of June 1 prior to gametime during which Robinson took batting practice and there is some instant of June 1 after gametime at which Baltimore will win.

It might be felt that this view sanctions an excessively lenient standard of truth. Suppose, for example, that Baltimore happened to be playing a double-header on June 1. We would be forced to accept the clause 'Baltimore will lose' as true if Baltimore will win the first game and lose the second. Surely, the critic might say, 'Baltimore will lose' is false -- or at best ambiguous -- when uttered in this context. Our reply is that the critic may well be right in calling a particular utterance of 'Baltimore will lose' ambiguous.

But it is ambiguous only because the possible reference times have not been made clear. Once the possible reference times are made clear we will certainly want to count the sentence true if it describes an event which occurs at any of these times. Suppose, for example, that our questionable clause had occurred in a paragraph beginning 'Next week will see many surprises in professional sports.' Surely we would then count our clause true if Baltimore should lose any game next week.

There is a second, related objection which might be raised to the treatment suggested here. A sentence like 'Baltimore won and Baltimore lost' sounds inconsistent; but on our interpretation there are contexts in which it comes out true. The sentence sounds inconsistent because ordinarily we infer 'It is not the case that Baltimore lost' from 'Baltimore won'. But this inference depends on two things: i) a tacit premise about the relation between winning and losing and ii) the assumption that premise and conclusion refer to the same time. Both of these can be made explicit in our formal framework, and we can then show that 'Baltimore won and Baltimore lost' is false.[4] It should be pointed out that a genuine contradiction (like 'Baltimore won and it is not the case that Baltimore won') is, on our account <u>always</u> false.

---

[4]In general it seems much more sensible to treat inferences of natural language by this two-pronged method of making premises explicit and finding semantical conditions corresponding to presuppositions than by merely adding premises until the argument goes through. Some presuppositions -- like the one that two clauses refer to the same time -- are clearly semantic.

B. The Language of ET (L_{ET})

It might be supposed that our language should contain one connective for each English tense. It turns out, however, to be more economical (and probably more illuminating) to take three tense connectives as primitive and show that any tensed sentence can be represented by applying these three to a group of basic formulas. Basic formulas are distinguished by the fact that their truth and appropriateness are independent of their 'utterance time'. Tense connectives, when applied to these, determine whether their reference time precedes or succeeds their time of utterance. It only makes sense to supply this information once, so tense connectives cannot be iterated. Since tense does not completely determine reference time, English contains a number of expressions which, when prefixed to tensed sentences, narrow further the range of possibilities. We include two of these in our formal language because their presence exposes some of the differences between the tenses.

Definition 3.1. The set Q of sorts of L_{ET} is the set of non-empty subsets of $\{u, t, I, I_s, <, >\}$.

Intuitively $\{u\}$, $\{t\}$, $\{I\}$, and $\{I_s\}$ are the one-dimensional sorts. $\{u\}$ is the sort of sentences whose truth depends only on utterance time. $\{t\}$, $\{I\}$, and $\{I_s\}$ are the sorts of sentences whose truth depends only on reference time: the reference time is either an instant, an interval or an interval-with-the-subinterval-property according to whether the sentence is of the first, second, or third of these sorts. A sentence of sort $\{<\}$ is one which refers to a time prior to its time of utterance, and a sentence of sort {>} is one which refers to a time after its time of utterance. Sentences of the other sorts are Boolean combinations of simpler sentences.

It is not always obvious which sort of formula should be used to represent an English sentence, but there are syntactic tests that can provide good clues. If it is appropriate to ask 'When?' after the utterance of a sentence, or to add a clause beginning 'when...' to the end of the sentence, then that sentence is not of sort u. If a sentence admits a progressive tense it must be either of sort I or sort I_s. If phrases like 'for a while' can be added to a sentence then the sentence is not of sort I. Facts like these suggest that our sort distinctions are more than a convenient theoretical device for making the truth conditions come out right. They play a role in grammar as well.

Definition 3.2. L_{ET} is the Q-sorted f-Boolean language[5] described below.

There are countably many sentence letters of sort t, sort I, and sort I_s. There are two unary basic connectives, '*Prg*' and '*Perf*'. '*Prg*' applies to sentences of sort {I} or sort $\{I_s\}$ to form a sentence of sort {t}. '*Perf*' applies to sentences of sort {t}, {I}, or $\{I_s\}$ to form sentences of sort {t}. There are three unary tense connectives, '*P*', '*F*' and '*N*'. '*P*' and '*F*'

---

[5]See Definition 2.5.

apply to sentences of sort $\{t\}$, $\{I\}$, or $\{I_s\}$ and yield sentences of sorts $\{<\}$ and $\{>\}$, respectively. '*N*' applies only to sentences of sort $\{t\}$ and yields sentences of sort $\{u\}$. Finally, there are two unary time-specifying connectives '*T*' and '*Y*' (for 'tomorrow' and 'yesterday' which apply to sentences of sorts $\{<\}$ and $\{>\}$ respectively, and yield sentences of sort $\{u\}$.

C. Models and truth

In this section we specify how the formal language L_{ET} is to be interpreted. Our initial formulations (Definitions 3.4 and 3.5) will be designed to conform as closely as possible to the general treatment for many-sorted systems presented in Chapter 2. We will then show (in Lemma 3.1) how the truth definitions can be written in a more perspecuous form.[6]

Definition 3.3. A time structure is a pair $(T,<)$ such that T is an infinite set and $<$ is a linear order on T. If $(T,<)$ is a time structure, we write '$a > b$' to mean $b < a$, '$a \geq b$' to mean not $a < b$, and '$a \leq b$' to mean not $b < a$. An interval on $(T,<)$ is a set $I \subseteq T$ such that if $u \in I$, $v \in I$, $u \leq x$ and $x \leq v$, then $x \in I$. If I is an interval on $(T,<)$ then Interior (I) = $\{x \in I : \exists y \in I, y > x \text{ and } \exists z \in I, z < x\}$.

---

[6]The reader who has difficulty understanding Definition 3.4 is urged to look ahead to Lemma 3.1.

A <u>temporal context in (T,<)</u> is a pair (a,R) such that $a \in T$ and $R \subseteq T \cup \{\text{intervals on } (T,<)\}$.

<u>Definition 3.4</u>. Let $(T,<)$ be a time structure and let I be the set of all intervals on $(T,<)$. An <u>ET-model on (T,<)</u> is an f-classical model suitable for L_{ET}[7] such that the following hold.

1) W is the set of all temporal contexts in $(T,<)$.
2) $(a,R)\ E_{\{u\}}\ (b,S)$ iff $a = b$.
3) $(a,R)\ E_{\{t\}}\ (b,S)$ iff $R \cap T = S \cap T$.
4) $(a,R)\ E_{\{I\}}\ (b,S)$ iff $R \cap I = S \cap I$.
5) $E_{\{I_s\}} = E_{\{I\}}$.
6) $(a,R)\ E_{\{<\}}\ (b,S)$ iff $\{x \in R\colon a > x\} = \{x \in S : b > x\}$.
7) $(a,R)\ E_{\{>\}}\ (b,S)$ iff $\{x \in R\colon a < x\} = \{x \in S : b < x\}$.
8) • $(a,R)\ \overline{Prg}(\{I\})\ (b,S)$ iff, for some X in I and some c in $R \cap T$, $S \cap I = \{X\}$ and $c \in \text{Interior}\ (X)$.
   • $\overline{Prg}\ (\{I_s\}) = \overline{Prg}\ (\{I\})$.
9) • $(a,R)\ \overline{Prg}\ (\{I\})\ (b,S)$ iff, for some $x \in T$ and some c in $R \cap T$, $S \cap T = \{x\}$ and $c > x$.
   • The clauses for $\overline{Perf}\ (\{I\})$ and $\overline{Perf}\ (\{I_s\})$ are the same except that '$x \in T$' is replaced by '$x \in I$' and '$S \cap T$' is replaced by '$S \cap I$'.
10) • $(a,R)\ \overline{P}\ (\{t\})\ (b,S)$ iff, for some x in $R \cap T$, $S \cap T = \{x\}$ and $a > x$.
   • The clauses for $\overline{P}\ (\{I\})$ and $\overline{P}\ (\{I_s\})$ are the same except that '$R \cap T$' is replaced by '$R \cap I$' and

---

[7] See Definitions 2.3b, 2.7.

'$S \cap T$', by '$S \cap I$'.

- The clauses for $\overline{F}(\{t\})$, $\overline{F}(\{I\})$, and $\overline{F}(\{I_s\})$ are the same as those for $\overline{P}(\{t\})$, $\overline{P(\{I\})}$ and $\overline{P(\{I_s\})}$, respectively, except that '$a > x$' is replaced by '$a < x$'.

11) $(a,R)\ \overline{N(\{t\})}\ (b,S)$ iff $S \cap T = \{a\}$.

12) $(a,R)\ F(>)\ (b,S)$ iff there is a c such that $S \cap \{x : b < x\} = \{c\}$ and $[a] + 1 < c$, $[a] + 2 > c$.

13) $(a,R)\ \overline{Y}(<)\ (b,S)$ iff there is a c such that $S \cap \{x : b > x\} = \{c\}$ and $[a] - 1 < c$, $[a] > c$.

14) If p is a sentence letter of sort $\{t\}$, $\{I\}$, or $\{I_s\}$, then $(a,R) \in V(p)$ iff $(a,\{b\}) \in V(p)$ for some $b \in R$. Furthermore, if p is of sort $\{I_s\}$, $(b, \{i\}) \in V(p)$ implies $(b, \{j\}) \in V(p)$ whenever j is a subinterval of i.

M is an <u>ET-model</u> if it is an ET-model on some time structure. It is easy to check that conditions 1-14 do determine a model suitable for L_{ET} in the sense of Definition 2.3b. For example, suppose $(a,R)\ \overline{F}(I)\ (b,S)$, $(b,S)\ E_I\ (b',S')$ and $(a,R)\ E_>\ (a',R')$. Then: 1) For some x in $R \cap I$, $S \cap I = \{x\}$ and $a < x$. 2) $S' \cap I = S \cap I$ and 3) $R \cap \{x : a < x\} = R' \cap \{x : a' < x\}$. By 1) and 2) $S' \cap I = \{x\}$. By 1) and 3), $a' < x$ and $x \in R' \cap I$. Hence $(a',R')\ F(I)\ (b',S')$.

Truth is defined a little differently than usual.

Definition 3.4. Let $M = (W, E, o, \bar{C}, V)$ be an ET-model and let (a,S) be in W. Then A is appropriate at (a,R) in M if one of the following holds.

- $A = PB$ and there is a b in R such that $a > b$.
- $A = FB$ and there is a b in R such that $a < b$.
- $A = Prg\ B$, $Perf\ B$, NB, TB or YB.
- A is a Boolean combination of formulas, all of which are appropriate at (a,R) in M.

A is true at (a,R) in M ($(a,R) \models A$) if A is appropriate at (a,R) and one of the following hold.

- $A = p$ and $(a,R) \in V(p)$.
- The main connective of A is Boolean and the usual conditions on the truth of A's main subformula (s) hold.

  $A = \square B$ (where $\square$ is non-Boolean, B is of sort-j, j is in the domain of the type of $\square$) and there is some (b,S) in W_M such that $(a,R)\ \bar{\square}(j)\ (b,S)$ and $(b,S) \models B$.

A is false at (a,R)

If a is appropriate at (a,R), but A is not true at (a,R).

A is appropriate (true, false) in M if it is appropriate (true, false) at the designated point in M.

Lemma 3.1. If M is an ET-model on $(T,<)$, I is the set of all intervals on $(T,<)$ and $(a,R) \in W_M$, then

1) $(a,R) \models Perf A$ iff $\exists b \in R \cap T$, $\exists c \in T \cup I$: $b > c$ $(a, \{c\}) \models A$.

2) $(a,R) \models Prg\ A$ iff $\exists b \in R \cap T$, $\exists c \in I$: $b \in$ Interior (c), $(a, \{c\}) \models A$.

3) $(a,R) \models PA$ iff $\exists s \in R$, $a > s$ $(a, \{s\}) \models A$.

4) $(a,R) \models FA$ iff $\exists s \in R$, $a < s$ $(a, \{s\}) \models A$.

D. The Tenses of English

The table on the next page indicates how English tenses are represented. Notice that not every formula of ET can represent an English sentence. In particular formulas of the form *Prg* p and *Perf*p do not by themselves correspond to English sentences. These formulas, together with the atomic ones, make up the class of basic formulas mentioned in section B. Tensed sentences are represented by applying the connectives '*P*', '*F*', and '*N*' to the basic formulas. In this section we list some of the important features of this representation.

1) Every English sentence not in the present tense is represented by an instant-evaluated formula. But the interval-evaluated formulas are still needed to account for the differences among the tensed sentences. [See 5, 6, 7 below.]

2) A sentence in the present tense is taken as saying that the event described takes place at some interval determined by context. So 'Dan drives to Detriot' means that a drive-to-

The Tenses Of English

Tense	English sentence	Representation in L_{ET}	Conditions for Truth at u With Reference to S in Model on $(T,<)$ with Valuation V
simple present	Sally swims. The cat is on the mat.	p q or Hq	$(u,S) \in V(p)$ $(u,S) \in V(q)$
simple past	Sally swam. The cat was on the mat.	Pp Pq	$\exists s\ (s \in S \cap I,\ u>s,\ (u, \{s\}) \in V(p))$ $\exists s\ (s \in S \cap T,\ u>s,\ (u, \{s\}) \in V(q))$
simple future	Sally will swim. The cat will be on the mat.	Fp Fq	$\exists s\ (s \in S \cap I,\ u<s,\ (u, \{s\}) \in V(p))$ $\exists s\ (s \in S \cap T,\ u<s,\ (u, \{s\}) \in V(q))$
present progressive	Sally is swimming.	$N\ Prg\ p$	$\exists t\ (t \in I,\ u \in Interior(t),\ (u, \{t\}) \in V(p))$
past progressive	Sally was swimming.	$P\ Prg\ p$	$\exists s\ \exists t\ (s \in S \cap T,\ t \subset I,\ u>s,\ s \in Interior(t),\ (u, \{t\}) \in V(p)$
future progressive	Sally will be swimming.	$F\ Prg\ p$	$\exists s\ \exists t\ (s \in S \cap T,\ t \in I,\ u<s,\ s \in Interior(t),\ (u, \{t\}) \in V(p)$
present perfect	Sally has swum. The cat has been on the mat.	$N\ Perf\ p$ $N\ Perf\ q$	$\exists t\ (t \in I,\ u>t,\ (u, \{t\}) \in V(p))$ $\exists t\ (t \in T,\ u>t,\ (u, \{t\}) \in V(q))$
past perfect	Sally had swum. The cat had been on the mat.	$P\ Perf\ p$ $P\ Perf\ q$	$\exists s\ \exists t\ (s \in S \cap T,\ t \in I,\ u>s,\ s>t,\ (u, \{t\}) \in V(p))$ $\exists s\ \exists t\ (s \in S \cap T,\ t \in T,\ u>s,\ s>t,\ (u, \{t\}) \in V(q))$
future perfect	Sally will have swum. *The cat will have been on the mat.*	$F\ Perf\ p$ $F\ Perf\ q$	$\exists s\ \exists t\ (s \in S \cap T,\ t \in I,\ u<s,\ s>t,\ (u, \{t\}) \in V(p))$ $\exists s\ \exists t\ (s \in S \cap T,\ t \in T,\ u<s,\ s>t,\ (u, \{t\}) \in V(q))$
present perfect progressive	Sally has been swimming.	$N\ Perf\ Prg\ p$	$\exists t\ \exists t'\ (t \in T,\ t' \in I,\ u>t,\ t \in Interior(t'),\ (u, \{t'\}) \in V(p))$
past perfect progressive	Sally had been swimming.	$P\ Perf\ Prg\ p$	$\exists s\ \exists t\ \exists t'\ (s \in S \cap T,\ t \in T,\ t' \in I,\ u>s,\ s>t,\ t \in Interior(t'),\ (u,\{t'\}) \in V(p))$
future perfect progressive	Sally will have been swimming.	$F\ Perf\ Prg\ p$	$\exists s\ \exists t\ \exists t'\ (s \in S \cap T,\ t \in T,\ t' \in I,\ u<s,\ s>t,\ t \in Interior(t'),\ (u,\{t'\}) \in V(p))$

Detroit-by-Dan takes place at the understood interval. Of course present tense English sentences are sometimes, (in fact usually) construed in other ways. But it can be argued that these other interpretations are derivative in the sense that they can be obtained by prefixing modifiers like 'occasionally' or 'customarily' to present tense sentences of the kind we are considering. One such 'derivative' interpretation gets special mention in the table -- present tense instant-evaluated sentences (like 'It is cloudy') are usually taken as saying that the event described takes place at the moment of utterance. Such sentences can be represented in ET by a formula of the form Nq.

3) The operators 'N', 'Y', and 'T' each supply the reference time at which succeeding clauses are to be evaluated. Hence none of these connectives can be applied to a formula which already contains an instance of any of them. So sentences like 'Yesterday John will come tomorrow' and 'Yesterday Yesterday it was Monday' are ungrammatical. 'Yesterday John came and tomorrow he will leave', however, is all right since there is no nesting of the operators.

4) 'John has built a house' is taken to mean that John is now a former builder-of-a-house. Since 'Yesterday' supplies a reference time for the sentence which follows it, 'Yesterday John had built' a house is well-formed, but 'Yesterday John has built a house' is not.

5) 'John built a house' is taken to mean that (at some prior interval supplied by context) 'John builds a house' was true. Hence, if John built a house two days ago, then 'Yesterday John built a house' is false, but 'Yesterday John had built a house' is true.

6) The progressive form of a sentence is true when it refers to an instant in the interior of an interval at which the original sentence is true. But all past and future tense sentences and many present tense sentences (like 'John owns a car') can only be true at <u>instants</u>. Hence none of these sentences have progressive forms. For the same reason, the operation of forming the progressive cannot be iterated.

7) 'John was building a house' is taken to mean that there was some instant contained in an interval at which 'John builds a house' is true. Hence 'John was building a house' is unlike 'John built a house', John 'has built a house' and 'John had built a house' in that it does not entail that a house was finished. (It does, however, entail that one will be finished eventually.)

8) To form the past or future tense of a basic sentence is to locate the event described in that sentence relative to the moment. This can be done at most once and, if the event is already located, there is no point in doing it at all. For this reason the operation of forming the past

tense isn't performed on sentences which are already in the past tense -- or on those which are in the future tense. In ET '*P*' and '*F*' don't apply to formulas beginning '*P*', '*F*' and '*N*', '*T*', and '*Y*'.

9) The distinctions among interval-evaluated sentences alluded to earlier are preserved in ET. 'If John is swimming in the Channel then he has swum in the Channel' is true in all ET-models, but 'if John is swimming across the channel then he has swum across it' is not.

10) No non-Boolean connectives can be applied to Boolean combinations of formulas of different sorts. If an English sentence appears to have this structure, it is being misread. Thus 'Professor Jones is writing his book and teaching' should not be represented by something of the form $N\ Prg\ (p \wedge q)$ but rather by a formula of the form $N\ Prg\ p \wedge N\ Prg\ q$. For otherwise we would be forced to call the sentence false unless the teaching began at least as early as the book writing. Similarly, 'Jones wrote his book and taught' should be represented by $P\ p \wedge P\ q$ rather than by $P(p \wedge q)$. These arguments do not apply to Boolean combinations of formulas of the same sort. $N\ Prg\ (q_1 \wedge q_2)$ is a well-formed formula of L_{ET} and may be taken as a translation of 'Jones is writing and teaching'. Notice, however, that this formula is equivalent to $N\ Prg\ q_1 \wedge N\ Prg\ q_2$ so this particular construction is not really needed. $N\ Prg\ (p_1 \wedge p_2)$ is also well-formed but it

is not equivalent to $N\ Prg\ p_1 \wedge N\ Prg\ p_2$. The question of which of these should be used to represent 'John is teaching Philosophy 150 and writing his book' is determined by context.

CHAPTER IV

QUANTIFIERS AS MANY-SORTED CONNECTIVES

A. Introduction

The observation that quantifiers and modal connectives resemble one another goes back at least as far as Von Wright's 1951 Essay in Modal Logic. Von Wright notes that "The logic of the words 'possible', 'impossible', and 'necessary',... is very similar to the logic of words 'some', 'no', and 'all', and attributes this similarity to the fact that "...the possible is that which is true under some circumstances, the impossible, that which is true under no circumstances." He calls the subject matter of quantification theory "the mode of existence" and includes this, along with the modes of "truth", "knowledge", and "obligation", in his list of major modal categories. Von Wright's aim was to exploit the similarity he observed by transferring well-known results and techniques of quantification theory to modal logics. Our aim is the reverse; we hope our knowledge of modal logic can be used to gain a better understanding of quantification theory. More specifically, we intend to construct many-sorted (propositional) modal logics with the same expressive power as first order predicate logic. First, however, we discuss briefly two earlier attempts along these lines -- the first by Arthur Prior, the second by Richard Montague.

Prior (in [Prior, 1968a] and [Prior, 1968b]) suggests a modal system called "egocentric logic" that looks a great deal like the ordinary monadic predicate calculus. Prior's reasoning goes something like this:[1] A distinctive feature of modal logic is that sentences are not considered to be true or false, but rather "true at a world", or "true at a time". Considered by itself a modal sentence is incomplete or "open", but given certain additional information it may become either true or false. But we can think of other examples of sentences whose truth depends in this way on extra information. Consider, for example, sentences whose subject is 'I'. At a given instant 'I am sitting' is true when uttered by some people and false when uttered by others. We find it fruitful to construct logics whose sentences are evaluated at worlds and times, so why not a logic whose sentences are evaluated at people? Or, stretching Prior's idea slightly, why not allow sentences to be evaluated at objects in general, so that 'I am inanimate' is a sentence true "at" the Washington Monument? Or (with a little more stretching) why not let some kinds of sentences by evaluated at pairs of objects so that 'My first member is the author of my second' is true at the pair (Scott, Waverly). Continuing in this manner we could allow sentences which are true at triples and quad-

---

[1]Actually Prior combines two arguments. First, that what is normally treated as a monadic predicate can be treated as a modal sentence to be evaluated at objects, and second, that what is normally treated as a binary predicate can be treated as the alternativeness relation of a Kripke model. We consider only the first argument.

ruples, or, more generally, we could allow sentences which are true at "assignments", or infintie sequences of objects. If modal logic can successfully deal with sentences which are open in the sense described, then we might expect it to be useful in dealing with open sentences of the more familiar kind, viz., sentences which are usually represented by a predicate letter followed by individual variables. A sentence which is true at all (some) worlds is necessarily (possibly) true; a sentence which is true at all (some) times is always (sometimes) true. Similarly a sentence true at all (some) assignments is universally (existentially) true. Thus, the most natural modal connectives correspond to quantifiers, and Prior's egocentric logic can be regarded as an attempt to modalize at least a part of first order predicate logic.

Montague, in one of his earliest papers,[2] points out many of the same similarities noted by von Wright. He observes, for example, that the theorems $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$, $\Box A \leftrightarrow \Box\Box A$, $\Box A \leftrightarrow (-\Box -A)$, of common modal systems are mirrored by the theorems $\forall x(A \rightarrow B) \rightarrow (\forall xA \rightarrow \forall xB)$, $\forall xA \leftrightarrow \forall x \forall xA$, $\exists xA \leftrightarrow (-\forall x-A)$ of quantification theory. These observations lead him to suggest the following uniform treatment of necessity and quantification. Let L be a language with predicate symbols, constant symbols, quantifiers, classical propositional connectives and a necessity operator $\Box$. Let a model M for this language by a triple (D,R,f) where $(D,R) = (D,(\overline{R}_1,\overline{R}_2,\ldots,\overline{c}_1,\overline{c}_2,\ldots))$ is an ordinary model for pre-

[2][Montague, 1960], which was written in 1955.

dicate logic and f is an assignment of elements of D to the variables of L. Let two models for L (D,R,f) and (D',R',f') be related by $\underline{P_\Box}$ if $D = D'$ and $f = f'$, by $\underline{P_x}$ if $D = D'$, $R = R'$, and f and f' assign the same element of D to every variable except possibly x. Now we can define the truth of a formula A in a model M so that the similarity between $\Box$ and $\forall$ surfaces:

> $\Box A$ is true in M if for all M', $M\ P\ M'$ implies A is true in M'.
>
> $\forall x\ A$ is true in M if for all M', $M\ P_x\ M'$ implies A is true in M'.

Thus the addition of quantifiers to a language is really no different than the addition of a collection of modal operators.

It would be natural, in light of the subsequent development of semantics for modal logic, to try to recast Montague's formulation in terms of relations between possible worlds, rather than relations between models. We could, of course, just let Montague's models _be_ the possible worlds. Then, if W is the class of all such models, o is a member of W, and V is a valuation which makes every atomic formula $R_i x_1 \ldots x_n$ true in exactly those worlds (D,R,f) such that $(f(x_1),\ldots,f(x_n))$ is in $\overline{R}_i$, then $(W, o, P, P_{v_1}, P_{v_2}, \ldots, V)$ is a Kripke model which verifies the same formulas as the Montague model o. ($v_1, v_2, \ldots$ is supposed to be an enumeration of all the individual variables.)

This solution, however, is not very satisfactory, for the class of Kripke models we obtain is very unnatural. Instead of writing out explicitly the truth values assigned by each V we could have said that we were interested in all models $(W, 0, P, P_{v_1}, \ldots, V)$ such that if (D,R,f) and (D,R,f') are in W and f assigns to x_i the same object as f' assigns to y_i for all i, $1 \leq i \leq n$, then for all predicate letters Q, $Qx_1 \ldots x_n$ is true at (D,R,f) under V if and only if it is true at (D,R,f') under V.

This account is better, but, if our object is to learn something about quantifiers by making use of what we know about modality, then it is not good enough. The logic determined by a class of models satisfying such a complicated condition is not likely to be a logic we know much about. The problem is that while Montague has, in a sense, replaced quantifiers, he has not replaced the predicates and variables that go with them. If $Qx_1 \ldots x_n$ and $Qy_1 \ldots y_n$ are treated as distinct sentence letters then (unless the class of possible valuation functions is restricted as above) the sentence $\exists x_1 \ldots \exists x_n\, Qx_1 \ldots x_n \rightarrow \exists y_1 \ldots \exists y_n\, Qy_1 \ldots y_n$ will be falsifiable. If, on the other hand, they are treated as the same sentence letter, then $\exists x_1 \ldots \exists x_n\; \exists y_1 \ldots \exists y_n\; (Px_1 \ldots x_n \wedge -Py_1 \ldots y_n)$ will not be satisfiable. In a logic without variables the formulas corresponding to $Px_1 \ldots x_n$ and $Py_1 \ldots y_n$ should not both be atoms.

The question of whether quantification theory can be formulated without variables and operators which bind them has received a great deal of attention.[3] The aim of the work that has been done is not to replace quantifiers by modal connectives. In some cases, however, we can twist it to suit this purpose. In particular, we consider in this chapter modal logical version (C and PF) of Tarski's cylindrical algebra ([Henkin, Monk, Tarski, 1974]) and of Quine's predicate functor logic ([Quine, 1960],[Quine, 1972]). Each of these logics is characterized by a class of Kripke models. The valuations in these models, like those in the models discussed above are not unrestricted. It turns out, however, that the restrictions needed are just those which mark our logics as <u>many-sorted</u>. Furthermore, the sorts play a familiar role. Each point of the models has a number of <u>dimensions</u> and the sort of a sentence determines which dimensions affect the truth value of that sentence.[4] In fact the logics we construct bear a strong resemblance to Krister Segerberg's two-dimensional logic B (which was formulated with a completely different interpretation in mind). To bring out this

---

[3]In addition to the works mentioned in this paragraph, see [Curry, Feys, 1958], [Schoenfinkel, 1924].

[4]Readers who are familiar with cylindric algebra might object that explicit sorting of sentence letters is not necessary, provided it is stipulated that for every sentence letter there is <u>some</u> n such that the truth of the sentence letter is unaffected by changes in dimensions beyond the n'th. This kind of move is legitimate in the case of cylindric algebra, because in that case we are only out to show that predicate logic has the structure of <u>some</u> cylindric algebra. In this chapter, however, we want to show that predicate logic is <u>equivalent</u> to cylindric and predicate functor logic. This will not necessarily be true under the stipulation mentioned. (For example, it will not be true if the truth of every sentence letter depends only on the first ten dimensions of a point.)

resemblance we construct a third logic B* which is a straightforward generalization of B. C, PF, and B* are all intended to be modal equivalents of the classical predicate calculus. In section B we describe these systems. In section C we investigate their relation to the predicate calculus, and in Section D, we consider the problem of axiomatization.

The material in this chapter supports the view that modal connectives (construed broadly enough) can do the work of quantifiers. Furthermore, it indicates that the modal way of thinking can shed new light on predicate-functor logic.

B. C, PF, and B*

1) C.

In Chapter I we remarked that the sentence letters and Boolean connectives of classical logics can be interpreted as the elements and operations of a Boolean algebra. But nothing was said about an algebraic interpretation of quantifiers. As a matter of fact, it is possible to generalize the notion of a Boolean-algebra in such a way to accommodate quantifiers. One generalization that will do the job is Tarski's ω-dimensional cylindric algebra (CA). The elements of an ω-dimensional cylindric algebra are sets of length-ω sequences of elements from some basic domain. There are the usual Boolean operations and constants (see p. 11), and, in addition there is a 'cylindrification operation' C_k for each natural number k and a 'diagonal constant', $D_{j,k}$ for each pair (j,k) of natural numbers. If X is a set, $C_k(X)$ is the set of all sequences which are obtained from members of X by changing at most the k'th coordinate. $D_{j,k}$ is the set of all sequences whose j'th and k'th coordinates are the same. To see the connection between CA and classical predicate calculus, suppose A is a sentence of a predicate language and suppose X is the set of all sequences $(d_1,d_2,\ldots)$ which "satisfy" A in the sense that A is true when d_1 is assigned to v_1, d_2 to v_2, and so forth. (Recall that $v_1,v_2,\ldots$ is an enumeration of all individual variables of our predicate language.) Then $C_k(X)$ is the set of all sequences which "satisfy" $\exists v_k A$. The diagonal elements are needed to express connections like that between Pv_1v_2 and Pv_2v_1.

Before describing our modal formulation C of CA, we introduce notation that will be useful throughout the remainder of the chapter.

$\underline{N^+}$ is the set of non-empty initial segments of the non-negative integers. $\underline{N} = N^+ \cup \phi$. We sometimes identify a non-negative integer with the set of all smaller non-negative integers. Thus N is the set of all non-negative integers and for all i, j in N, $i \cup j = \max(i,j)$. For all k in N, $\underline{r_k} : N \to N$ is defined by $r_k(j') = j$ if $j \neq k$ and $r_k(k) = k-1$ if $k \neq 0$, $r_0(0) = 0$. i is the identity function on N. Finally, if X is a set of sequences, $\underline{E^*_k(X)} = \{(d,e) : d \in X,\ e \in X$ and for all i such that $1 \leq i \leq k$, $d_i = e_i\}$, $E^*_0(X) = X^2$ and $E^*(X) = \{E^*_k(X) : k \in N\}$.

<u>Definition 4.1a</u>. L_c is the N-sorted f-Boolean propositional language with:

(1) sort-j sentence letters $P^j_1, P^j_2, \ldots$ for all j in N

(2) connectives '$\tilde{C}_k$' of type r_k for all k in N and 0-ary connectives 'd_{ij}' of type $i \cup j$ for all i,j in N.
'C_k' is used as an abbreviation for '$- \tilde{C}_k -$'.[5]

<u>Definition 4.1b</u>. If D is a non-empty set, a <u>C-model</u> on D is an f-classical Kripke-model $(W, E, o, \bar{\tilde{c}}_1, \ldots, \bar{d}_{1,1}, \ldots, V)$ suitable for L_c such that

(1) $W = D^\omega$.

(2) $E = E^*(W)$.

(3) $d \; \bar{\tilde{c}}_k(n) \; e$ iff $d_i = e_i$ for all i, $1 \leq i \leq n$ such that $i \neq k$.

Notice that, for any non-empty D, there is a C-model on D. If A is a sentence of sort $\{1, \ldots, n\}$ then the truth of A at a sequence w depends only on the first n coordinates of w.

<u>Notation</u>: If $\mathcal{M}$ be the class of all C-models, then $\models_c = \models_{\mathcal{M}}$ and C (the <u>cylindric logic</u>) = $\mathcal{L}(\models_c)$.

2) PF.

Quine's <u>predicate functor logic</u>, unlike CA, was devised

---

[5] We take '$\tilde{C}_k$' as primitive rather than 'C_k' in order to insure that all our connectives are Kripke.

specifically as a solution to the problem of formulating the classical predicate calculus without variables and variable binding operators. His language contains the familiar n-ary predicate letters, and in addition, a finite number of functors which operate on predicates to yield new ones. For example, '$\cap$' and 'p' are functors such that, if P and Q are binary predicates, then $\cap$PQ and pQ are also binary predicates. We think of $\cap$PQ as the predicate which holds of a pair (x,y) iff both P and Q hold of (x,y), and pQ as the predicate which holds of (x,y) if Q holds of (y,x). The language of predicate functor logic, however, does not contain individual variables. Hence we cannot write down the obvious axioms which would express these properties. To get the effect of existential quantification there is a functor ']' such that]P holds of $(x_1,\ldots,x_n)$ iff for some x_0, P holds of $(x_0,x_1,\ldots,x_n)$. (Notice that if P is an n-ary predicate,]P is an n-1-ary predicate.)

Our modal version of PF is very similar to PF itself. Several of Quine's functors (including '$\cap$') become ordinary f-Boolean connectives and we prefer to deal with]'s dual instead of] itself. Quine, of course, did not think of his predicate functors as Kripke connectives nor of his predicates as modal sentences. But our semantics is clearly not far from what he had in mind. Furthermore, our modal treatment enables us to describe a proof procedure for PF, a task which for Quine was still "a major agendum."

<u>Definition 4.2a</u>. L_{PF} is the N^+-sorted f-Boolean propositional language such that:

(1) The sort-j sentence letters are those of L_C.

(2) The non-Boolean connectives are '[', '$\tilde{]}$', 'p', 'P' and 'I', where '$\wedge$', '$\vee$', '$\rightarrow$' are type-m; 'p', 'P' and '-' are type-i, '[' is type r^+ where $r^+(j) = j+1$ for all j in ω. '$\tilde{]}$' is type r^- where $r^-(j) = j-1$ if $j \in N^+$, $r^-(0) = 0$; and I is a 0-ary connective of type 2. ']' is used as an abbreviation for '-$\tilde{]}$-'.

<u>Definition 4.2b</u>. If D is a non-empty set, a PF-model on D is an f-classical Kripke model $(W, E, o, \overline{[}, \overline{\tilde{]}}, \overline{P}, \overline{p}, \overline{I}, V)$ suitable for L_{PF} such that:

(1) $W = D^\omega$.

(2) $E = E^*$.

(3) $\vec{d}\ \overline{[}(n)\ \vec{e}$ iff for all $i \leq n$, $e_i = d_{i+1}$.

(4) $\vec{d}\ \overline{\tilde{]}}(n)\ \vec{e}$ iff $\vec{e}\ \overline{[}(n)\ \vec{d}$.

(5) $\vec{d}\ \overline{P}(n)\ \vec{e}$ iff $e_i = d_n$ and for $2 \leq i \leq n$, $e_i = d_i-1$.

(6) $\vec{d}\ \overline{p}(n)\ \vec{e}$ iff $\vec{d}\ \overline{P}(2)\ \vec{e}$.

(7) $\overline{I} = \{\vec{d} \in W : d_1 = d_2\}$.

<u>Notation</u>: If M is the class of all PF models, then $\underline{\models_{PF}} = \models_M$ and $\underline{PF} = L(\models_{PF})$.

3) B*.

Segerberg's logic B (for 'basic') was devised to provide a solution to a problem of Åqvist's in the field of tense logic.[6] In addition to the Boolean connectives it contains the unary connectives '$\Box$', '$\boxed{\vert}$', '$\boxminus$', '$\obar$', '$\ominus$', and '$\otimes$'. Segerberg suggests that these connectives be read as 'everywhere', 'everywhere on this longitude', 'everywhere on this latitude', 'at the diagonal point of this lognitude', 'at the diagonal point of this latitude', and 'at the mirror image point', respectively. From the truth conditions he gives, however, it is clear that another reading is possible. We can think of the sentences of B as formulas of a predicate language with only two variables, say x and y. If A is interpreted $C(x,y)$, $\Box A$ becomes $\forall x \forall y\, C(x,y)$, $\boxed{\vert}A$ becomes $\forall y\, A(x,y)$, $\boxminus A$ becomes $\forall x\, A(x,y)$, $\obar A$ becomes $A(x,x)$, $\ominus A$ becomes $A(y,y)$ and $\otimes A$ becomes $A(y,x)$. Thus B can be interpreted as a fragment of predicate logic. To get all of predicate logic we need a way to express formulas with more than 2 or less than 2 variables. Toward this end we divide the sentence letters of B into infinitely many sorts. We also need connectives which express quantification with respect to variables other than 'x' and 'y'. Toward this end we rename $\boxminus$ by $\boxed{1}$, $\boxed{\vert}$ by $\boxed{2}$ and add the connectives $\boxed{3}$, $\boxed{4}$, ... with similar truth conditions involving coordinates beyond the second. ($\boxed{k}$

---

[6] See [Åqvist, 1973].

turns out to be the same as $\tilde{C}_k$ of L_C.) We retain ①, but ⊖ and □ are both dispensable (□A is $\boxed{1}\ldots\boxed{n}$A where $n = s(A)$ and ⊖A is ①ⓧA). Finally, to insure that we can get the effect of arbitrary permutations of variables we replace ⓧ by ⓧ and Ⓧ (which are the same as p and P of L_{PF}). The details are presented below.

Definition 4.3a. L_{B*} is the N-sorted f-Boolean language with

(1) The same sentence letters as L_C.

(2) The type-i connectives Ⓧ, ⓧ, and ① .

(3) For all k in N, a type-r_k connective $\boxed{k}$.

Definition 4.4b. A B*-model on D is an f-classical Kripke model (W, E, o, $\overline{\boxed{1}}$, $\overline{\boxed{2}}$, ..., $\overline{①}$, $\overline{Ⓧ}$, $\overline{ⓧ}$, V) suitable for L_{B*} such that:

(1) $W = D^{\omega}$.

(2) $E = E*$.

(3) $\overline{\boxed{k}} = \bar{\tilde{C}}_k$, $\overline{Ⓧ} = \overline{P}$, $\overline{ⓧ} = \overline{p}$ (see Definitions 4.3b, 4.2b).

(4) $\vec{d}\ \overline{①}\ \vec{e}$ iff $e_1 = e_2 = d_1$, and for $i > 2$, $e_i = d_i$.

Notation: If M is the set of all B*-models, let $\models_{B*} = \models_M$ and $B* = L(\vdash_{B*})$.

Notation: The interpretations for the connectives in C-models, PF-models and D-models are uniquely determined from the points in those models. Hence if it is known that M is a PF-model (C-model, B*-model) we represent M by (W_M, O_M, V_M). If M is known from context and $w \in W_M$, then we write '$w \models A$' for $(M,w) \models A$.

C. Equivalence with Predicate Logic

A prominent difference between C and PF on the one hand, and B* on the other, is that the former have sentence constants (i.e., 0-ary connectives), while the latter does not. The work that is done by the unary connective '①' in B* is done in PF by the constant 'I' and in C by the constants 'd_{ij}'. This difference, it turns out, is not merely a matter of style. For I and the d_{ij}'s take us beyond Pred C, while ① does not. More specifically, I and the d_{ij}'s give the effect of adding the special binary predicate, '=' ('equality') to the language of Pred C and endowing it with the expected properties. Since this modification of Pred C was not discussed in earlier chapters we make a short digression to do so now.

<u>Definition 4.5a</u>. $\underline{L_{PCE}}$ is the (one-sorted) predicate language with no connectives except the Boolean ones and the quantifiers and with the special binary predicate letter '='. '=xy' is normally written 'x=y'. For all n we assume that $P^n_1, P^n_2, \ldots$ is an enumeration of the n-ary predicate letters of L_{PCE}.

<u>Definition 4.5b</u>. $\underline{\vdash_{PCE}}$ is the smallest classical consequence relation in L_{PCE} satisfying:

(1) $\phi \vdash X = X$.

(2) $X = Y \vdash Y = X$.

(3) $X = Y \quad Y = Z \vdash X = Z$.

(4) $A(x,x_1,\ldots,x_n),\ x = y \vdash A(y,x_1,\ldots,x_n)$ for all formulas A.

<u>Definition 4.5c</u>. A <u>PCE-model</u> is a classical model $M = (D, \overline{P}, a)$ suitable for L (see p. 18) such that $\overline{=_M}$ is $\{(d,d) : d \in D\}$. $\underline{\mathcal{M}_{PCE}}$ is the class of all PCE-models.

<u>Theorem 4.1</u>. If $\mathcal{M} = \mathcal{M}_{PCE}$ then $\models_{\mathcal{M}} = \vdash_{PCE}$.

We omit the proof.

<u>Theorem 4.2</u>.

a) $\vdash_C \approx \vdash_{PCE}$.

b) $\vdash_{PF} \approx \vdash_{PCE}$.

c) $\vdash_{B*} \approx \vdash_{Pred\ C}$.

C, PF, and B* are alike in that sentences are evaluated at sequences rather than assignments. In proving Theorem 4.2 we make use of a natural correspondence between sequence talk and assignment talk. g_1^* is a function from C-models to PCE-models (or from PF-models to PCE-models or from B*-models to Pred C models) such that if $M = (D^{\omega}, o, V)$, $g_1^*(M) = (C, \overline{P}, a)$ where $C = D$, $\overline{P_j^n} = \{(d_1, \ldots, d_n)$: for some $\vec{e}$ in D^{ω} , $\vec{e} \in V(P_j^n)$ and $(d_1, \ldots, d_n)$ is an initial segment of $\vec{e}\}$, and $a(v_i)$ is the i'th coordinate of O. $\underline{g_2^*}$ is a function in the other direction such that, if $M = (C, \overline{P}, a)$, $g_2^*(M) = (D^{\omega}, o, V)$ where $D = C$, $o = (a(v_1), a(v_2), \ldots)$ and $V(p_j^n) = \{(d_1, d_2, \ldots) : (d_1, \ldots, d_n) \in \overline{P_j^n}\}$. It will always be clear from the context what kind of models the arguments and values of g_1^* and g_2^* are intended to be. Notice that $g_2^*(g_1^*(M)) = g_1^*(g_2^*(M)) = M$.

To prove the theorem, we must find functions from our special propositional languages to the ordinary predicate languages and back, which correspond to these translations of models. Clearly our functions should take the sentence letters P_j^n to $P_j^n v_1 \ldots v_n$. But n-ary predicate letters of L_{PCE} and $L_{Pred\ C}$ can also be followed by other strings of variables. We will have to show that, in each of our propositional languages, the effect of relettering the variables in $P_j^n v_1 \ldots v_n$ can be obtained by applying connectives to P_j^n.

Notation. If A is a sort-n sentence of L_C and $k_1, \ldots, k_n$ are positive integers, then '$\tau_{(k_1, \ldots, k_n)} A$' denotes the sentence $c_{m+1} \ldots c_{m+n}\ (d_{k_1, m+1} \wedge \ldots \wedge d_{k_n, m+n} \wedge c_1 \ldots c_n$ $(A \wedge d_{1, m+1} \wedge \ldots \wedge d_{n, m+n})$.

Property. if $\vec{d} = (d_1, d_2, \ldots)$ and $\vec{e} = (e_1, e_2, \ldots)$ are members of W such that $e_1 = d_{k_1}, \ldots, e_n = d_{k_n}$ and A is a sort-n sentence of L_C, then $\vec{d} \models \tau_{\langle k_1,\ldots,k_n \rangle} A$ iff $\vec{e} \models A$.

Proof of a. Let $f_2\colon L_{PCE} \to L_C$ be defined as follows.

- If $A = P_j^0$, $f_2(A) = A$.
- If $A = P_j^{n+1} v_{k_1} \ldots v_{k_{n+1}}$, $f_2(A) = \tau_{\langle k_1,\ldots,k_{n+1} \rangle} P_j^{n+1}$.
- If $A = v_k = v_j$, $f_2(A) = d_{k,j}$.
- If $A = -B$, $f_2(A) = -f_2(B)$.
- If $A = B \square C$ where $\square$ is '$\wedge$', '$\vee$', or '$\to$', then $f_2(A) = f_2(B) \square f_2(C)$.
- If $A = \forall v_k B$, $f_2(A) = \breve{C}_k\ f_2(B)$.
- If $A = \exists v_k B$, $f_2(A) = C_k\ f_2(B)$.

Claim: For all A in L_{PCE} and all PCE-models M, $M \models A$ iff $g_2^*(M) \models t_2(A)$.

Proof. A routine induction on A using the property of $\tau_{\langle k_1,\ldots,k_n \rangle}$ mentioned above.

The other direction is easy. Let $f_1\colon L_C \to L_{PCE}$ be defined as follows:

- $f_1(p_j^0) = P_j^0$.
- $f_1(p_j^{n+1}) = P_j^{n+1} v_1 \ldots v_{n+1}$.
- $f_1(d_{j,k}) = v_j = v_k$.
- $f_1(B \square C) = f_1(B) \square f_1(C)$ if $\square$ is '$\wedge$', '$\vee$' or '$\to$'.
- $f_1(-A) = -f_1(A)$.
- $f_1(\breve{C}_k A) = \forall v_k\ t_1(A)$.

<u>Claim</u>: For all A in L_C and all C-models M, $M \models A$ iff $g_1^*(M) \models f_1(A)$.

<u>Proof</u>. Another routine induction on A.

From the corollary to Lemma 1.23 it now follows that $\vdash_C \approx \vdash_{PCE}$.

<u>Proof of b</u>. Instead of showing that $\models_{PF}$ is strongly similar to $\models_{PCE}$ directly, it is convenient to show that they are both strongly similar to an intermediate system. In L_{PF} both variables and variable-binding operators are done away with in favor of predicate functors. In the intermediate language L_{PFV} we keep the variables, but eliminate the binding operators.

<u>Definition 4.6a</u>. The alphabet of L_{PFV} consists of <u>atomic predicates</u> which we take to be the predicates of L_{PCE} <u>predicate functors</u>, which we take to be the connectives of L_{PF}, and <u>individual variables</u>, which we take to be the individual variables of L_{PCE}. The class of <u>n-ary predicates of L_{PFV}</u> is defined inductively: All n-ary predicates of L_{PCE} are n-ary predicates of L_{PFV}. If $P = \Box P_1 \ldots P_n$ where $\Box$ is a type-r connective of L_{PF}, $P_1, \ldots, P_n$ are $j_1, \ldots, j_n$-ary predicates of L_{PFV}, respectively, and $r(j_1, \ldots, j_n) = n$, then P is an n-ary predicate of L_{PFV}. <u>L_{PFV}</u> is the set of all strings of the form $Px_1 \ldots x_n$ such that P is an n-ary predicate of L_{PFV} and $x_1, \ldots, x_n$ are individual variables. We can easily give an interpretation of L_{PFV} in terms of models for L_{PCE}.

Definition 4.6b. If M is a PCE-model and A is in L_{PFV}, then A is true in M (written: $M \models A$)[7] if one of the following conditions holds:

1. $A = P_j^0$ or $A = P_j^n x_1 \ldots x_n$ and $M \models A$.

2. $A = Ixy$ and $M \models x = y$.

3. $A = P\, P^n x_1 \ldots x_n$ and $M \models P^n x_n x_1 \ldots x_{n-1}$.

4. $A = p\, P^n x_1 \ldots x_n$ and $M \models P^n x_2 x_1 x_3 \ldots x_n$.

5. $A = [\ P^{n-1}\ x_1 \ldots x_n$ and $M \models P^{n-1}\ x_2 \ldots x_n$.

6. $A = \tilde{]}\ P^{n+1}\ x_1 \ldots x_n$ and, for some x_0 not among $x_1, \ldots, x_n$, $M' \models P^{n+1} x_0 \ldots x_n$ for all $M' = \langle D_M, \overline{P}_M, a' \rangle$ such that $a'(x) = a(x)$ whenever $x \neq x_0$.

7. $A = (R^m \wedge P^n) x_1 \ldots x_{\max(m,n)}$ and $M \models R^m x_1 \ldots x_m$, $M \models P^m x_1 \ldots x_n$.

8. $A = (R^n \vee P^n) x_1 \ldots x_{\max(m,n)}$ and $M \models R^m x_1 \ldots x_m$ or $M \models P^n x_1 \ldots x_n$.

9. $A = (R^m \rightarrow P^n) x_1 \ldots x_{\max(m,n)}$ and either not $M \models R^m x_1 \ldots x_n$ or $M \models P^n x_1 \ldots x_n$.

10. $A = -P^n x_1 \ldots x_n$ and not $M \models P^n x_1 \ldots x_n$.

An easy induction establishes that if $M = (D, \overline{P}, a)$, $M' = (D, \overline{P}, a')$, and for $1 \leq i \leq n$, $a'(x_i) = a(v_i)$, then

---

[7] We are safe in using the word 'truth' and the symbol '$\models$' because the two definitions for $M \models A$ coincide when $A \in L_{PCE} \cap L_{PFV}$.

$M \models P^n\, x_1 \ldots x_n$ iff $M \models P^n\, v_1 \ldots v_n$.

If $\Gamma \subseteq L_{PFV}$ and $A \in L_{PFV}$ let $\Gamma \models_{PFV} A$ iff every A is true in any PCE-model in which all the members of Γ are true. It is easy to verify that $\models_{PFV}$ is a consequence relation.

We now show that $\models_{PFV} \underset{\vee}{\approx} \models_{PCE}$.

The function $f_1\colon L_{PFV} \to L_{PCE}$ is defined by induction.

- $f_1(Ixy) = x = y$.
- $f_1(P^n_j\, x_1 \ldots x_n) = P^n_j\, x_1 \ldots x_n$.
- $f_1(PP^n\, x_1 \ldots x_n) = f_1(P^n\, x_n x_1 \ldots x_{n-1})$.
- $f_1(pP^n\, x_1 \ldots x_n) = f_1(P^n\, x_2 x_3 \ldots x_n)$.
- $f_1([P^{n-1}\, x_1 \ldots x_n) = f_1(P^{n-1}\, x_2 \ldots x_n)$.
- $f_1(\tilde{]}\, P^{n+1}\, x_1 \ldots x_n) = \forall x_0 f_1(P^{n+1}\, x_0 x_1 \ldots x_n)$ where x_0 is the first variable not among $x_1, \ldots, x_n$.
- If $\Box =$ '$\wedge$', '$\vee$' or '$\to$', $f_1((R^m \Box P^n)x_1 \ldots x_{\max(m,n)}) = f_1(R^n\, x_1 \ldots x_m) \Box f_1(P^n\, x_1 \ldots x_n)$.
- $f_1(-P^n\, x_1 \ldots x_n) = -f_1(P^n\, x_1 \ldots x_n)$.

<u>Claim 1</u>: For all A in L_{PFV}, $M \models A$ iff $M \models f_1(A)$.

<u>Proof</u>. By induction on the number of functor-occurrences in A. We consider the case that $A =]\, P^{n+1} x_1 \ldots x_n$. Then $M \models A$ iff for all $M' = (D, \overline{P}, a')$ such that $a'(x) = a(x)$ whenever $x \neq x_0$, $M' \models P^{n+1}\, x_0 \ldots x_n$. By induction hypothesis this is equivalent to

$M' \models f_1(P^{n+1}\, x_0 \ldots x_n)$ for all such M'; i.e., to $M \models \forall x_0\, f_1(P^{n+1}\, x_0 \ldots x_n)$. But $\forall x_0\, f_1(P^{n+1}\, x_0 \ldots x_n)$ is $f_1(A)$.

To define our translation in the other direction it is convenient to introduce some notation.

Definition 4.7a.[8] if $1 \leq k \leq n$ and P^n is an n-ary predicate of L_{PFV}, let

1) $\sigma_k\, P^n \;=\; P^{(n-k)+1}\, (Pp)^{k-1}\, P^n$.

2) $\sigma_k^{-1}\, P^n \;=\; \sigma_k^{k-1}\, P^n$.

3) $S_k P^n \;=\;]\, (I \wedge \sigma_k^{-1} P^n)$.

σ_k has the effect of moving the k'th variable in a list to the front, σ_k^{-1} has the effect of moving the first variable in a list to the k'th place and S_k has the effect of adding a second occurrence of the first variable in a list in the k'th place, in a sense which is made precise by the following lemma.

Lemma 4.1a.

1) $M \models \sigma_k P^n x_1 \ldots x_n$ iff $M \models P^n x_k x_1 \ldots x_{k-1} x_{k+1} \ldots x_n$.

2) $M \models \sigma_k^{-1} P^n x_1 \ldots x_n$ iff $M \models P^n x_2 \ldots x_k x_1 x_{k+1} \ldots x_n$.

3) $M \models S_k P^n x_1 \ldots x_{n-1}$ iff $M \models P^n x_1 \ldots x_{k-1} x_1 x_k \ldots x_{n-1}$.

4) $f_1(\sigma_k P^n x_1 \ldots x_n) = P^n x_k x_1 \ldots x_{k-1} x_{k+1} \ldots x_n$.

---

[8] Here and in what follows positive superscripts indicate iterations. Thus '$\Box^2 A$' means $\Box\Box A$, and '$(\Box\bigcirc)^2 A$' means $\Box\bigcirc\Box\bigcirc A$.

5) $f_1(\sigma_k^{-1} P^n x_1 \ldots x_n) \;=\; P^n x_2 \ldots x_k x_1 x_{k+1} \ldots x_n.$

6) $f_1(S_k P^n x_1 \ldots x_{n-1}) \;=\; -\; x_0 - (x_0 = x_1 \wedge f_1(P^n x_1 \ldots x_{k-1}).$

(Of course, $x_{k+1} \ldots x_n$ and $x_k \ldots x_{n-1}$ in the right hand side of 1-6 are empty if $k = n$.)

The proofs of 1-6 are straightforward.

Now we define $f_2\colon L_{PCE} \rightarrow L_{PFV}$ by induction on the number of connective occurrences in the argument.

- $f_2(P_j^n\, x_1 \ldots x_n) \;=\; P_j^n\, x_1 \ldots x_n.$
- $f_2(x = y) \;=\; Ixy.$

Suppose f_2 is defined on sentences of L_{PCE} with k or fewer connective-occurrences and that A and B are such sentences and $f_2(A) = P^n\, x_1 \ldots x_n$, $f_2(B) = R^m\, x_{n+1} \ldots x_{n+m}$.

- $f_2(-A) \;=\; -P\, x_1 \ldots x_n.$
- If $\square$ is '$\wedge$', '$\vee$', or '$\rightarrow$', $f_2(A \,\square\, B) \;=$ $(P^n \,\square\, [^n\, R^m)\; x_1 \ldots x_{n+m}.$
- If x is not among $x_1, \ldots, x_n$ then $f_2(\exists xA) =$ $f_2(\forall xA) = f_2(A)$. Otherwise let $k_1, \ldots, k_p$ be a list (in increasing order) of all the i's such that $x_i = x$, and let $(z_1, \ldots, z_{n-p})$ be the sequence which results from $(x_1, \ldots, x_n)$ by deleting the x_{k_i}'s. Then
  $f_2(\exists xA) \;=\;](S_{k_2} \ldots S_{k_p}\, \sigma_k^{-1}\, P^n)\; z_1 \ldots z_{n-p}$ and
  $f_2(\forall xA) \;=\; \tilde{]}(S_{k_2} \ldots S_{k_p}\, \sigma_{k_1}^{-1}\, P^n)\; z_1 \ldots z_{n-p}.$

Notice that if A is any formula of PCE, the free variables of A are included in the variables of $f_2(A)$.

Claim 2. For all formulas A of PCE and all PCE-models M, $M \models A$ iff $M \models f_2(A)$.

Proof. By induction on the length of A. We do only two cases.

- Suppose $A = B \wedge C$, $f_2(B) = P^n x_1 \ldots x_n$, $f_2(C) = R^m y_1 \ldots y_m$, $M \models B$ iff $M \models P^n x_1 \ldots x_n$ and $M \models C$ iff $M \models R^m y_1 \ldots y_m$. Then $M \models A$ iff $M \models P^n x_1 \ldots x_n$ and $M \models R^m y_1 \ldots y_m$. This holds iff $M \models P^n x_1 \ldots x_n$ and $M \models [{}^n R^m x_1 \ldots x_n y_1 \ldots y_m$, i.e., iff $M \models (P^n \wedge [{}^n R^m)\ x_1 \ldots x_n y_1 \ldots y_m$ $(= f_2(A))$.

- Now suppose $A = \exists x B$, $f_2(B)$ is as above and $M \models B$ iff $M \models f_2(B)$. then there are two subcases:
  i) x is not among $x_1, \ldots, x_n$. In this case $M \models A$ iff $M \models B$ iff $M \models f_2(B)$ $(= f_2(A))$.
  ii) For some $k_1, \ldots, k_p$, $(p \leq n)$ $x_{k_1} = x_{k_2} = \ldots = x_{k_p} = x$. Let $z_1, \ldots, z_{n-p}$ be the sequence obtained from $x_1, \ldots, x_n$ by deleting the x_{k_j}'s and suppose that $k_1 < \ldots < k_p$. Then $M \models f_2(A)$ iff, for some $M = (\mathcal{D}_M, \mathcal{O}_M, a')$ such that a' agrees with a_M except possibly on x, $M' = S_{k_2} \ldots S_{k_p} \sigma_{k_1}^{-1} P^n x\, z_1 \ldots z_{n-p}$ (since x is not among $z_1, \ldots, z_{n-p}$). By Lemma 4.1a, this holds iff there is such an M' such that $M' \models P^n x_1 \ldots x_n$. By induction hypothesis this holds iff there is such an M' such that $M' \models B$, i.e., iff $M = \exists x B$ $(=A)$.

By the corollary to Lemma 1.23, Claims 1 and 2 establish that $\models_{PFV} = \models_{PCE}$.

We now work toward proving that $\models_{PFV} \approx \models_{PF}$.

Definition 4.7b. If A^n is a sort-n sentence of L_{PF}

1) $\sigma_k A^n = P^{(n+1)-k}\ (Pp)^{k-1}\ A^n$ if $k \leq n$

$= \sigma_k (A^n\ [^{k-2} - \perp))$ otherwise.

2) $\sigma_k^{-1} A^n = \sigma_k^{k-1} A^n$.

3) $\tau_k A^n = \rceil\ \sigma_{k+1}(I \wedge \sigma_{k+1}^{-1}\ A^n)$.

4) $\tau_{\langle k_1,\ldots,k_n\rangle}\ A^n = \tau_{k_n}\ \tau_{k_{n-1}+1} \cdots \tau_{k_1+n-1}\ A^n$.

Lemma 4.1b. If $A^n \in L_{PF}$ and A^n is sort-n,

1) $(d_1,\ldots,d_k,\ldots) \models \sigma_k A^n$ iff $(d_k,d_1,\ldots,d_{k-1},d_{k+1},\ldots) \models A^n$.

2) $(d_1,\ldots,d_k,\ldots) \models \sigma_k^{-1} A^n$ iff $(d_2,\ldots,d_k,d_1,d_{k+1},\ldots) \models A^n$.

3) $(d_1,\ldots,d_k,\ldots) \models \tau_k A^n$ iff $(d_k,d_1,d_2,\ldots) \models A^n$.

4) $(d_1,\ldots,d_k,\ldots) \models \tau_{\langle k_1,\ldots,k_n\rangle} A^n$ iff $(d_{k_1},\ldots,d_{k_n},d_1,d_2,\ldots) \models A^n$.

(The notation $w \models A$ was explained on p.164). Notice that if A is of sort-n, $\sigma_k A^n$ is of sort max(n,k), $\tau_k A^n$ is of sort max(n-1,k) and $\tau_{\langle k_1,\ldots,k_n\rangle} A$ is of sort $\max(k_1+n-1,\ k_2+n-2,\ldots,k_n)$.

To define the function from L_{PFV} to L_{PF}, recall that every predicate of L_{PFV} is a sentence L_{PF}. If $A = P^0$, let $f_2(A) = A$. If $A = P^{n+1}\ v_{k_1} \ldots v_{k_{n+1}}$, let $f_2(A) = \tau_{\langle k_1,\ldots,k_{n+1}\rangle} P^{n+1}$.

Claim 3a. For all PCE-models M, and all A in L_{PFV}: $M \models A$ iff $g_2^*(M) \models f_2(A)$.

Proof. It is easy to prove by induction on the complexity of P^n that $M \models P^n v_1 \ldots v_n$ iff $g_2^*(M) \models P^n$. The claim follows by Lemma 4.1b and the definition of f_2.

Define $f_1\colon L_{PF} \to L_{PFV}$ by $f_1(P^0) = P^0$, $f_1(P^{n+1}) = P^{n+1} v_1 \ldots v_{n+1}$.

Claim 3b. For all PF-models M and all A in L_{PF}: $M \models A$ iff $g_1^*(M) \models f_1(A)$.

Proof. $g_1^*(M) \models f_1(A)$ iff $M\ (= g_2^*(g_1^*(M))) \models f_2 f_1(A)$ by 3a. But $f_2(f_1(A))$ is equivalent to A by Lemma 4.1b, whence the claim follows.

By the corollary to Lemma 1.23, claims 3a and 3b establish that $\models_{PCE} \approx \models_{PF}$.

Proof of c. As before, it is convenient to introduce some defined connectives.

Definition 4.7c. If A^n is a sort-n sentence of L_{B*}

1) $\sigma_k A^n = ⓧ^{(n+1)-k} (ⓧ ⓧ)^{k-1} A^n$ if $k \leq n$ and $\sigma_k (A^n \wedge (p^k \vee -p^k)$ otherwise.

2) $\sigma_k^{-1} A^n = \sigma_k^{k-1} A^n$.

3) $\tau_k A^n = \sigma_{n+1} \sigma_{k+1} ① \sigma_{k+1}^{-1} A^n$.

4) $\tau_{\langle k_1, \ldots, k_n \rangle} A^n = \tau_{k_n} \tau_{k_{n-1}+1} \cdots \tau_{k_1+n-1} A^n$.

Lemma 4.1c. 1-4 of Lemma 4.1b hold when A^n is a sort-n sentence of L_{B*}.

f_2: Pred C $\rightarrow L_{B*}$ is defined as follows:

- If $A = P^0_j$, $f_2(A) = A$.
- If $A = P^n_j v_{k_1} \dots v_{k_n}$, $f_2(A) = \tau_{\langle k_1,\dots,k_n \rangle} P^n_j$.
- If $A = B \square C$ where $\square$ is '$\wedge$', '$\vee$', or '$\rightarrow$' then $f_2(A) = f_2(B) \square f_2(C)$.
- If $A = -B$, $f_2(A) = -f_2(B)$.
- If $A = \forall v_k B$, $f_2(A) = \boxed{k}\, f_2(B)$.
- If $A = \exists v_k B$, $f_2(A) = -\boxed{k} -f_2(B)$.

Claim 1. For all A in $L_{\text{Pred C}}$ and all Pred-C-models M, $M \models A$ iff $g^*_2(M) \models f_2(A)$.

The proof of this claim is similar to proofs of the previous claims.

f_1: $L_{B*} \rightarrow$ Pred C is defined by induction:

- $f_1(P^n_j) = P^n_j v_1 \dots v_n$.
- $f_1(Ⓧ A)$ is the result of interchanging all occurrences of v_n and v_1 in $f_1(A)$ where n is the sort of A.
- $f_1(ⓧ A)$ is the result of interchanging all occurrences of v_1 and v_2 in $f_1(A)$.
- $f_1(① A)$ is the result of replacing all occurrences of v_2 by v_1 in $f_1(A)$.

- $f_1(\boxed{k}A) = \forall v_k\ f_1(A)$.
- $f_1(A \mathbin{\square} B) = f_1(A) \mathbin{\square} f_1(B)$ if $\square$ is '$\wedge$', '$\vee$', or '$\rightarrow$'.
- $f_1(-A) = -f_1(A)$.

Claim 2. For all A in L_{B*} and all B*-models M, $M \models A$ iff $g_1^*(M) \models f_1(A)$.

Proof. Induction on A.

By the corollary to Lemma 1.23, Claims 1 and 2 establish that $\models_{B*} \approx \models_{PCE}$.

D. Axiomatizations

Our aim in this section will be to axiomatize $\models_C$, $\models_{PF}$, and $\models_{B*}$. There are, of course, many axiomatizations of Pred C and PCE. The representation theorem for cylindric algebras (discussed, for example, in [Henkin, Tarski, 1961]) provides an 'axiomatization' of cylindric algebras. And there is a rather complicated completeness proof for B* in [Segerberg, 1973]. It would probably be possible to adapt any of these results to the frameworks we are considering and to use translation theorems like those of the preceding section to extend them to PF. We choose not to do this, however, for two reasons. First, having used the idea of a many-sorted modal logic to give a unified treatment of the three systems in question, it would be desirable to show that the completeness arguments of Chapters I and II can be modified to suit our purposes here. Second, since our motive for studying these systems was to add to our knowledge of the older systems, it seems ill-advised to base our axiomatizations upon the ones already given.

What we shall do, therefore, is to list certain principles which are valid for the interpretation we have in mind and prove that this list is sufficient. No attempt will be made to find the most economical set of axioms or even an independent one. We begin with some definitions.

<u>Definition 4.8</u>. If A is in L_C (L_{PF}, L_{B*}) then A <u>depends on coordinate k</u> if $k \geq 1$, $A \neq \perp$, and one of the following holds:

- $A = p_j^n$ and $1 \leq k \leq n$.
- $A = B \wedge C$, $A = B \vee C$, or $A = B \rightarrow C$, and either B or C depends on coordinate k.
- $A = -B$ and B depends on coordinate k.
- $A = \tilde{c}_j B$ or $A = \boxed{j} B$ and B depends on k and $k \neq j$.
- $A = d_{ij}$ and $k = i$ or $k = j$.
- $A = [B$ and B depends on coordinate $k-1$.
- $A =]B$ and B depends on coordinate $k+1$.
- $A = I$ and $k = 1$ or $k = 2$.
- $A = \perp B$ and B depends on k and $k \neq 2$.

- $A = \mathrm{P}\, B^n$ or $A = \textcircled{X} B^n$ and either $k = 1$ and B^n depends on coordinate n or $2 \leq k \leq n$ and B^n depends on coordinate $k-1$.
- $A = \mathrm{p}\, B^n$ or $A = \textcircled{x} B^n$, $n \geq 2$, and either $k = 1$ and B^n depends on coordinate 2 or $k = 2$ and B^n depends on coordinate 1.
- $A = \mathrm{P}B^1$, $\textcircled{X} B^1$, $\mathrm{p}B^1$, or $\otimes B^1$ and B^1 depends on coordinate k.

The terminology here is suggested by the following easily proved facts.

Lemma 4.2. a) If A does not depend on coordinate k then for all d_k, d_k^1, $(d_1,\ldots,d_{k-1}, d_k, d_{k+1},\ldots) \models A$ iff $(d_1,\ldots,d_{k-1}, d_k^1, d_{k+1},\ldots) \models A$.

b) If $\tau_{\langle k_1,\ldots,k_n \rangle} A^n$ depends on j then $j \in \{k_1,\ldots,k_n\}$.

Definition 4.9. If A is a sort-n sentence of L_C, L_{PF}, or L_{B*} then $A^* = \tau_{\langle 2,4,\ldots,2n \rangle} A$. If Γ is a set of sentences, then $\Gamma^* = \{A^* : A \in \Gamma\}$.

Axioms and rules (where n is the sort of A and $i \leq n$)

(τ1) $\tau_{\langle 1,\ldots,n \rangle} A \dashv\vdash A$.

(τ2) $\tau_{\langle k_1,\ldots,k_{n+m} \rangle} A \dashv\vdash \tau_{\langle k_1,\ldots,k_{n+p} \rangle} A$.

(τ3) $-\tau_{\langle k_1,\ldots,k_n \rangle} A \dashv\vdash \tau_{\langle k_1,\ldots,k_n \rangle} -A$.

(τ4) $\tau_{\langle k_1\ldots k_n \rangle} A \wedge \tau_{\langle k_1\ldots k_n \rangle} B \dashv\vdash \tau_{\langle k_1\ldots k_n \rangle} (A \wedge B)$

(C1) $\Gamma, \tau_{\langle k_1,\ldots,k_n \rangle} A \vdash \Delta$ implies

$\Gamma,\ \tau_{\langle k_1,\ldots,k_{i-1},m,k_{i+1},\ldots,k_n\rangle} C_i A \vdash \Delta$ (provided no sentence in the conclusion of this rule depends on k_j).

(C2) $\phi \vdash d_{j,j}$.

(C3) $d_{j,k} \vdash d_{k,j}$.

(C4) $d_{j,k},\ d_{km},\ \vdash d_{j,m}$.

(C5) $\tau_{\langle k_1,\ldots,k_n\rangle} A,\ d_{k_i,j} \vdash \tau_{\langle k_1,\ldots,k_{i-1},j,k_{i+1},\ldots k_n\rangle} A$.

(C6) $d_{k_i,k_j} \dashv\vdash \tau_{\langle k_1,\ldots,k_n\rangle} d_{i,j}$ provided $n \geq \max(i,j)$.

(C7) $\tau_{\langle k_1,\ldots,k_n\rangle} A \vdash \tau_{\langle k_1,\ldots,k_{i-1},k,k_{i+1},\ldots k_n\rangle} C_i A$.

(PF1) $\Gamma,\ \tau_{\langle k_0,\ldots,k_n\rangle} A \vdash \Delta$ implies

$\Gamma,\ \tau_{\langle k_1,\ldots,k_n\rangle}]A \vdash \Delta$ (provided no sentence in the conclusion of this rule depends on k_0).

(PF2) $\phi \vdash \tau_{\langle i,i\rangle} I$.

(PF3) $\tau_{\langle i,j\rangle} I \vdash \tau_{\langle j,i\rangle} I$.

(PF4) $\tau_{\langle i,j\rangle} I,\quad \tau_{\langle j,k\rangle} I \vdash \tau_{\langle i,k\rangle} I$.

(PF5) $\tau_{\langle k_1,\ldots,k_n\rangle} A,\quad d_{k_i,j} \vdash \tau_{\langle k_1,\ldots,k_{i-1},j,k_{i+1},\ldots,k_n\rangle} A$.

(PF6) $\tau_{\langle k_1,\ldots,k_n\rangle} PA \dashv\vdash \tau_{\langle k_n,k_1,\ldots,k_{n-1}\rangle} A$.

(PF7) $\tau_{\langle k_1,\ldots,k_n\rangle} pA \dashv\vdash \tau_{\langle k_2,k_1,k_3,\ldots,k_n\rangle} A$.

(PF8) $\tau_{\langle k_1,\ldots,k_n\rangle} [A \dashv\vdash \tau_{\langle k_2,\ldots,k_n\rangle} A$.

(PF9) $\tau_{\langle k_0,\ldots,k_n\rangle} A \vdash \tau_{\langle k_1,\ldots,k_n\rangle}]A.$

(B*2) $\sigma_i\sigma_j \textcircled{I} \sigma_j^{-1}\sigma_i^{-1} A,\ \sigma_j\sigma_k \textcircled{I} \sigma_k^{-1}\sigma_j^{-1} A \vdash \sigma_i\sigma_k \textcircled{I} \sigma_k^{-1}\sigma_i^{-1} A$

whenever $i < j < k$.

(B*3) $\sigma_i\sigma_j \textcircled{I} \sigma_j^{-1}\sigma_i^{-1} A,\ \sigma_k\sigma_j \textcircled{I} \sigma_j^{-1}\sigma_k^{-1} A \vdash \sigma_k\sigma_i \textcircled{I} \sigma_i^{-1}\sigma_k^{-1} A$

whenever $k < i < j$.

(B*4) $\sigma_i\sigma_j \textcircled{I} \sigma_j^{-1}\sigma_i^{-1} A,\ \sigma_k\sigma_j \textcircled{I} \sigma_j^{-1}\sigma_k^{-1} A \vdash \sigma_i\sigma_k \textcircled{I} \sigma_j^{-1}\sigma_i^{-1} A$

whenever $i < k < j$.

(B*5) $A \vdash \sigma_{k_i}\sigma_j \textcircled{I} \sigma_j^{-1}\sigma_{k_i}^{-1} A$ implies $\tau_{\langle k_1,\ldots,k_n\rangle} A \vdash$

$\tau_{\langle k_1,\ldots,k_{i-1},j,k_{i+1},\ldots,k_n\rangle} A.$

(B*1) $\Gamma,\ \tau_{\langle k_1,\ldots,k_n\rangle} A \vdash \Delta$ implies

$\Gamma,\ \tau_{\langle k_1,\ldots,k_n\rangle} - \boxed{j} - A \vdash \Delta$ (provided no sentence in the

conclusion of this rule depends on k_j).

(B*6) $\tau_{\langle k_1,k_2,\ldots,k_n\rangle} \textcircled{I} A \dashv\vdash \tau_{\langle k_1,k_1,k_3,\ldots,k_n\rangle} A.$

(B*7) $\tau_{\langle k_1,\ldots,k_n\rangle} \textcircled{X} A \dashv\vdash \tau_{\langle k_n,k_1,\ldots,k_{n-1}\rangle} A.$

(B*8) $\tau_{\langle k_1,\ldots,k_n\rangle} \textcircled{x} A \dashv\vdash \tau_{\langle k_2,k_1,\ldots,k_n\rangle} A.$

(B*9) $\tau_{\langle k_1,\ldots,k_j,\ldots,k_n\rangle} A \vdash \tau_{\langle k_1,\ldots,k_j,\ldots,k_n\rangle} - \boxed{j} - A.$

Definition 4.10. $\vdash_C$ is the smallest classical consequence relation on L_C satisfying $\tau 1$ - $\tau 4$ and C1 - C7. $\vdash_{PF}$ is the smallest classical consequence relation on L_{PF} satisfying $\tau 1$ - $\tau 4$ and PF1 - PF9. $\vdash_{B*}$ is the smallest classical consequence relation on L_{B*} satisfying $\tau 1$ - $\tau 4$ and B*1 - B*9.

Theorem 4.3. $\vdash_C = \models_C$, $\vdash_{PF} = \models_{PF}$ and $\vdash_{B*} = \models_{B*}$.

To prove soundness (i.e., that $\vdash_C \subseteq \models_C$, $\vdash_{PF} \subseteq \models_{PF}$, and $\vdash_{B*} \subseteq \models_{B*}$) is routine. To prove sufficiency, assume $\Gamma \nvdash_C \Delta$ ($\Gamma \nvdash_{PF} \Delta$, $\Gamma \nvdash_{B*} \Delta$). We must construct a model M such that $M \models \Gamma$ but, for all $B \in \Delta$, $M \not\models B$.

Definition 4.11. A theory (Γ,Δ) of L_C is $\vdash_C$-saturated if it is maximal $\vdash_C$ consistent and if $\tau_{\langle k_1,\ldots,k_j,\ldots,k_n \rangle} c_j A \in \Gamma$ implies that, for some k_j', $\tau_{\langle k_1,\ldots,k_j',\ldots,k_n \rangle} A \in \Gamma$. A theory (Γ,Δ) of L_{PF} is $\vdash_{PF}$-saturated if it is maximal $\vdash_{PF}$-consistent and if $\tau_{\langle k_1,\ldots,k_n \rangle}]A \in \Gamma$ implies that, for some k_0, $\tau_{\langle k_0,\ldots,k_n \rangle} A \in \Gamma$. A theory (Γ,Δ) of L_{B*} is $\vdash_{B*}$-saturated if it is maximal $\vdash_{B*}$-consistent and if $\tau_{\langle k_1,\ldots,k_j,\ldots,k_n \rangle} - \boxed{j} - A \in \Gamma$ implies that, for some k_j', $\tau_{\langle k_1,\ldots,k_j',\ldots,k_n \rangle} A \in \Gamma$.

Lemma 4.2. For all subsets Γ and Δ of L_C (L_{PF}, L_{B*}), if (Γ^*, Δ^*) is $\vdash_C$ ($\vdash_{PF}$, $\vdash_{B*}$)-consistent it can be extended to a $\vdash_C$ ($\vdash_{PF}$, $\vdash_{B*}$)-saturated theory (Γ^+, Δ^+).

Proof. We prove the lemma for Γ and Δ subsets of L_C. The other cases are proved similarly. Let $A_0, A_1, A_2, \ldots$ be an enumeration of L_C such that if $A_i = \tau_{\langle k_1,\ldots,k_n\rangle} C_i B$ and $i \leq n$ then $A_{i+1} = \tau_{\langle k_1,\ldots,k_{i-1},m,k_{i+1},\ldots k_n\rangle} B$. For some m such that m is an odd number, $m \notin \{k_1,\ldots,k_n\}$ and, for all $k < i$, A_k does not depend on m (such an m will always be available because each A_k depends on at most finitely many coordinates). We define a sequence of theories (Γ_i,Δ_i) as follows.

$$(\Gamma_0,\Delta_0) = (\Gamma^*,\Delta^*)$$

$$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{A_n\} & \text{if } (\Gamma_n \cup \{A_n\}, \Delta_n) \text{ is } \vdash_C\text{-consistent} \\ \Gamma_n & \text{otherwise.} \end{cases}$$

$$\Delta_{n+1} = \begin{cases} \Delta_n & \text{if } (\Gamma_n \cup \{A_n\}, \Delta_n) \text{ is } \vdash_C\text{-consistent} \\ \Delta_n \cup \{A_n\} & \text{otherwise.} \end{cases}$$

Let $\Gamma^+ = \bigcup_{i<\omega} \Gamma_i$ and $\Delta^+ = \bigcup_{i<\omega} \Delta_i$.

Note that all the axioms for $\vdash_C$ are virtually finite and all the rules preserve virtual finiteness. Hence $\vdash_C$ is finitary (see p. 23). From this it follows easily that (Γ^+,Δ^+) is maximal $\vdash_C$ consistent (see Lemma 1.9). All that remains is to show that (Γ^+,Δ^+) is $\vdash_C$-saturated. Suppose $\tau_{\langle k_1,\ldots,k_n\rangle} c_j A$ is A_n. Then if $A_n \in \Gamma^+$, A_n must be in Γ_{n+1}. We claim that $A_{n+1} = \tau_{\langle k_1,\ldots,k_{j-1},m,k_{j+1},\ldots,k_n\rangle} A$ is in Γ_{n+2}. If this were false, then we would have $\Gamma_{n+1}, A_{n+1} \vdash \Delta_{n+1}$. But m is an odd number, so nothing in $\Gamma^* \cup \Delta^*$ depends on it, and it was also chosen so that none of the A_j's listed before A_{n+1} depend on it. So $\Gamma_{n+1} \cup \Delta_{n+1}$ does not depend on m. Furthermore, m is not among $k_1, \ldots, k_n$, so A_{n+1} does not depend on m. By C1, therefore, $\Gamma_{n+1} A_n \vdash \Delta_{n+1}$, which violates the consistency of $(\Gamma_{n+1}, \Delta_{n+1})$. Hence (Γ^*, Δ^*) is $\vdash_C$-saturated.

Lemma 4.3. If $\Gamma^* \nvDash \Delta^*$ $(\Gamma^* \nvDash_{PF} \Delta^*,\ \Gamma^* \nvDash_{B^*} \Delta^*)$ then $\Gamma^* \nvDash_C \Delta^*$ $(\Gamma^* \nvDash_{PF} \Delta^*,\ \Gamma^* \nvDash_{B^*} \Delta^*)$.

Proof. If the hypothesis is true, then (Γ^*, Δ^*) is consistent, so by the previous lemma it can be extended to a saturated theory $t^+ = (\Gamma^+, \Delta^+)$. We will use this theory to build a model which shows $\Gamma \nvDash \Delta^*$. Consider the following binary relations on positive integers

$$i \sim_C j \quad \text{iff} \quad d_{ij} \in t^+$$

$$i \sim_{PF} j \quad \text{iff} \quad \tau_{\langle ij \rangle} I \in t^+$$

$$i \sim_{B^*} j \quad \text{iff} \quad \begin{array}{l} i = j \quad \text{or} \\ i < j \quad \text{and, for all } A \text{ in } L_{B^*},\ A \in t^+ \\ \qquad\qquad \text{implies } \sigma_i \sigma_j \; \textcircled{I} \; \sigma_j^{-1} \sigma_i^{-1} A \in t^+ \\ j < i \quad \text{and } j \sim_{B^*} i \end{array}$$

From axioms (C2-C4 (PF2-PF4, B*2-B*4) it follows that $\sim_C$ $(\sim_{PF},\ \sim_{B^*})$ is an equivalence relation. Let $[i]$ be the equivalence class of i under this relation. Let $D = \{[i] : i$ is a positive integer$\}$, and let $O = ([1], [2], \dots)$. Finally, let $V(P^n_j) = \{([k_1], [k_2], \dots) : \tau_{\langle k_1,\dots,k_n \rangle} P^n_j \in t^+\}$. Notice that whether a sequence belongs to $V(P^n_j)$ depends only on the first n places of the sequence. Hence $M = (D^\omega, o, V)$ is a C (PF, B*)-model.

<u>Claim</u>. For all sort-n sentences A in L_C (L_{PF}, L_{B*}), $([k_1],[k_2],\dots) \models A$ iff $\tau_{\langle k_1,\dots,k_n \rangle} A \in t^+$.

<u>Proof</u>. By induction on the length of A. A few cases will serve to illustrate:

• If A is a sentence letter then $([k_1],[k_2],\dots) \models A$ iff for some $j_1,\dots,j_n$, $[j_1] = [k_1],\dots, [j_n] = [k_n]$ and $\tau_{\langle j_1,\dots,j_n \rangle} A \in t^+$. By Axiom C5 (axiom PF5, rule B*5) this happens iff $\tau_{\langle k_1,\dots,k_n \rangle} A \in t^+$.

• If $A = d_{ij}$ then $([k_1],[k_2],\dots) \models A$ iff $[k_i] = [k_j]$ iff $d_{k_j,k_j} \in t^+$. By C6 this holds if $\tau_{\langle k_1,\dots,k_n \rangle} d_{ij} \in t^+$.

• If $A = \check{]}B$ where B is sort $n+1$, then $\tau_{\langle k_1,\dots,k_n \rangle} A \in t^+$ implies (by $\tau 3$) that $\tau_{\langle k_1,\dots,k_n \rangle}] - B \notin t^+$. This, in turn, implies that for all k_0, $\tau_{\langle k_0,\dots,k_n \rangle} - B \notin t^+$ (by PF9). Re-applying $\tau 3$, $\tau_{\langle k_0,\dots,k_n \rangle} B \in t^+$ for all k_0. By induction hypothesis, then, $([k_0],[k_1],\dots,[k_n]) = B$ for all k_0; i.e., $([k_1],\dots,[k_n]) =]B$. Conversely, if $\tau_{\langle k_1,\dots,k_n \rangle}]B \notin t^+$, then $\tau_{\langle k_1,\ldots,k_n \rangle}] - B \in t^+$. Since t^+ is saturated this means $\tau_{\langle k_0,\ldots,k_n \rangle} - B \in t^+$ for some k_0. Hence $\tau_{\langle k_0,\ldots,k_n \rangle} B \notin t^+$ and, by induction hypothesis, $([k_0],\ldots[k_n]) \nVdash B$; i.e., $([k],\ldots,[k_n]) \nVdash \tilde{]}B$.

From the claim, it follows that $([1],[2],\ldots) \models A^n$ iff $\tau_{\langle 1,2,\ldots n \rangle} A^n \in t^+$, iff (by $\tau 1$) $A^n \in t^+$. Hence $M \models A$ iff $A \in t^+$. This proves the lemma.

<u>Lemma 4.4</u>. If (Γ,Δ) is consistent so is (Γ^*,Δ^*).

We prove this lemma for PF-consistency. The other systems can be handled similarly. Suppose $\Gamma^* \vdash_{PF} \Delta^*$. Since $\vdash_{PF}$ is finitary, $A_1,\ldots,A_m \vdash_{PF} B_1,\ldots, B_n$ for some $A_1,\ldots,A_n$ and $B_1,\ldots,B_n$ in Γ^* and Δ^*, respectively. We know each A_i is of the form $\tau_{\langle 2,4,\ldots,2s_i \rangle} A_i^-$ and each B_j is of the form $\tau_{\langle 2,4,\ldots 2t_j \rangle} B_j^-$. By $\tau 2$ and $\tau 4$, then $\tau_{\langle 2,4,\ldots 2k \rangle}((A_1 \wedge \ldots \wedge A_m) \to (B_1 \vee \ldots \vee B_n)) \vdash \phi$ where $k = \max(\{s_i : 1 \leq i \leq m\} \cdot \{t_j : 1 \leq j \leq n\})$. By k applications of rule PF1, $]^k((A_1 \wedge \ldots \wedge A_n) \to (B_1 \vee \ldots \vee B_n)) \vdash \phi$. By k applications of Axiom PF9, $\tau_{\langle 1,\ldots,k \rangle}((A_1 \;\ldots\; A_m) \to (B_1 \;\ldots\; B_n)) \vdash \phi$. By $\tau 1$, $(A_1 \wedge \ldots \wedge A_m) \vdash (B_1 \vee \ldots \vee B_n)$ which violates the consistency of (Γ,Δ).

To complete the proof of Theorem 4.3, suppose $\Gamma \nvdash \Delta$. By Lemma 4.4 , $\Gamma^* \nvdash \Delta^*$. By Lemma 4.3 there is a model

$M = (D^{\omega}, (d_1, d_2, \ldots), V)$ such that $M \models \Gamma^*$ and, for all C in Δ^*, $M \not\models C$. Let $M^* = (D^{\omega}, (d_2, d_4, \ldots)\ V)$. Clearly $M^* \models A$ iff $M \models A^*$. So $M^* \models \Gamma$ and, for all D in Δ, $M^* \not\models D$.

Remarks on the Literature

In Chapter II we can distinguish two ideas, each of which has been discussed in recent literature. First, there is the syntactic point that a formal language in which every connective applies to every sentence is a poor imitation of natural language. It is better to rule the awkward strings ungrammatical than to force an interpretation on them. Second, there is the semantic point that the truth-in-a-situation of different sentences may depend on different features of the situation (and that dependence on inappropriate features is a common cause of the failure of a connective to apply to certain sentences).

The syntactic point was made by Arthur Prior in Appendix C of [Prior 1957]. In fact, Prior considers a restricted language for necessity which is the same as our L^*. He calls $S5 \cap L^*$ the system 'A', and guesses at an axiomatization for it. From our characterization of $\vdash^*_{S5}$, it is easy to check that Prior's guess is correct. (This also follows from the work in [Pollock, 1967].)

The semantic point (that logics should be capable of dealing with sentences whose truth values depend on different features of the situation) was made by Dana Scott in [Scott 1968]. Scott recommends a kind of model similar to those introduced in Sections C and D of Chapter 2 here. His models have only one kind of point, but each point has several coordinates, each of which stores a different kind of information about the situation. Richard Montague in [Montague 1968] discusses the same models as an example of the kind of system that can be subsumed within his very general framework for pragmatics. As an illustration he outlines a semantics for languages to deal simultaneously with tenses and pronouns in which each point in a model has a speaker-coordinate and a time-coordinate. Other examples have cropped up in more recent writings. Krister Segerberg ([Segerberg 1973]) calls the logics determined by this kind of model "n-dimensional" (where n is the number of coordinates of the points) and proves a completeness theorem for the "basic" two-dimensional logic B which is discussed here in Chapter 4. Hans Kamp ([Kamp 1971]) discusses two-dimensional systems for dealing with constructions involving words like "Now", but points out that the two-dimensional treatment in this case is merely a convenience. Frank Vlach ([Vlach 1973]) considers certain natural extensions of Kamp's logics in which the two dimensions play an essential role. Finally, David Kaplan ([Kaplan 1973]) considers a logic for demonstratives which has (among other special features) four dimensions. [Kaplan 1973] also contains a discussion of a problem analagous to our "incompatibility of choice sets".

It is easy to see that any logic which we call many-sorted is many-dimensional. For we can identify a point in a D-model with the sequence $(W/E_{j_1},\ldots,W/E_{j_n})$ where $(j_1,\ldots,j_n)$ is an enumeration of the sorts. The converse of this, however, is not true; for the interpretation of connectives and sentence letters in a many-dimensional model may be unrestricted. Roughly speaking, a many-dimensional, but single-sorted logic is one in which the truth of a sentence depends on several features of the situation, but the truth of every sentence depends on all these features. For example, consider a language with both the traditional tense operator 'F' (read "It will be the case that...") and the traditional possibility operator '$\Diamond$' (read "possibly..."). If we believe that what is possible later might be different from what is possible now (for example, if we take 'possible' to be 'technologically possible'), and if we believe that what comes later might possibly not come later, then every sentence in our language would depend on two coordinates -- a time coordinate and a world coordinate We would then then have a two dimensional, single-sorted logic. If we believe that possibility is not time-dependent (for example, if we take 'possible' to be 'logically possible,' then it would be more plausible to include in our language sentences of two sorts: world-dependent sentences and world-and-time-dependent sentences. The connective '$\Diamond$' would apply to either sort of sentence and the result would always be a world-dependent sentence. We would then have a two-sorted logic with one "two-dimensional sort" and one "one-dimensional sort". The combination of several sorts with several dimensions illustrated by this example also occurs in the systems discussed in Chapters III and IV.

The treatment of tenses in Chapter III shares many features with various treatments in the literature. The distinction between past and present perfect is brought out in [Reichenbach 1947]. The relation between a sentence and its progressive form is discussed as an example in [Scott 1968]. Many of the other points here are made in [McCawley 1971], although truth conditions for the tenses are not given explicitly here. Finally, arguments showing the weakness of the traditional Priorian tense logics as representations of English tenses have recently been given in [Needham 1975]. Needham, however, opts for a system in which quantification over times is allowed in the object language.

BIBLIOGRAPHY

[Åqvist 1973] Lennart Åqvist. "Modal Logic with Subjunctive conditionals." Journal of Philosophical Logic 2. (1973): pp. 2-76.

[Curry, Feys 1958] H.B. Curry and Robert Feys. Combinatory Logic. Vol. 1. Amsterdam: North Holland Publishing Co., 1958.

[Fine 1974] Kit Fine. "An incomplete logic containing K4." Theoria 60 (1974): pp. 23-30.

[Gabbay 1974a] Dov Gabbay. "Tense Logics and the Tenses of English." In Readings in Logic and Philosophy, Ed. by J.M.E. Moravcsik. The Hague: Mouton & Co., 1974.

[Gabbay 1974b] ________. "A Generalization of the Concept of Intensional Semantics." Philosophia 4 (1974): pp. 251-269.

[Gabbay f] ________. Semantical Methods in Non-Classical Logics. Vol. 1: Modal and Tense Logics. Dordrecht, D. Reidel Publishing Company, forthcoming.

[Henkin, Tarski 1961]. Leon Henkin and Alfred Tarski. "Cylindric Algebras." In Proceedings of Symposia in Pure Mathematics Vol. 2: Lattice Theory Providence: American Mathematical Society, pp. 82-113.

[Henkin, Monk, Tarski]. Leon Henkin, J. Donald Monk, and Alfred Tarski. Cylindric Algebras. Amsterdam: North Holland Publishing Co., 1971.

[Hughes, Cresswell 1968] G.E. Hughes and M.J. Cresswell. An Introduction to Modal Logic, London: Methuen and Co., 1968; University Paperbacks, 1972.

[Kamp 1971] Hans Kamp. "Formal Properties of 'Now'", Theoria 37 (1971), pp. 237-273.

[Kanger 1968] Stig Kanger. "Equivalent Theories", Theoria 34 (1968): pp. 1-6.

[Kaplan 1973] David Kaplan. "The Logic of Demonstratives". Los Angeles, 1971, expanded 1973 (Mineographed).

[Kleene 1952] Stephen Kleene. Introduction to Metamathematics. Princeton: D. Van Nostrand Company, 1952.

[Kotas, Pieczkowski 1970] Jerzy Kotas and August Pieczkowski. "Allgemeine Logische und Mathematische Theorien". Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik 16 (1970): pp. 353-376.

[McCawley 1971] James McCawley. "Tense and Time Reference in English". In Studies in Linguistic Semantics. New York: Holt, Rinehart and Winston, 1971.

[Montague 1960] Richard Montague. "Logical Necessity, Physical Necessity Ethics and Quantifiers". Inquiry 4 (1960): pp. 259-269. Reprinted in Formal Philosophy, Ed. by Thomason. New Haven: Yale University Press, 1974: pp. 71-84.

[Montague 1962] __________. "Pragmatics". In Contemporary Philosophy, Ed. by Klibansky. La Nuova Italia Editrice, 1968, pp. 102-122.

[Needham 1975] Paul Needham, Temporal Perspective: A Logical Analysis of Temporal Reference in English. No. 25: Philosophical Studies. Uppsala: Philosophical Society and the Department of Philosophy of the University of Uppsala, 1975.

[Ohnishi, Matsumoto 1961] Masao Ohnishi and Kazuo Matsumoto. "Gentzen Method in Modal Calculi." Osaka Mathematical Journal, 13, (1961): pp. 125-137.

[Penner 1970] Terence Penner. "Verbs and the Identity of Actions -- A Philosophical Exercise in the Interpretation of Aristotle," in Ryle -- A Collection of Critical Essays, Ed. by Wood and Pitcher. New York: Doubleday, 1970.

[Pollock 1967] John Pollock. "Basic Modal Logic." Journal of Symbolic Logic 32 (1967): pp. 355-365.

[Prior 1957] Arthur Prior. Time and Modality. London: Oxford University Press, 1957.

[Prior 1968a] ___________. "Egocentric Logic." Nous 2 (Aug. 1968): pp. 191-207.

[Prior 1968b] ___________. "Quasi-Propositions and Quasi Individuals." In Papers on Time and Tense, Ed. by Prior. London: Oxford University Press, 1968.

[Quine 1960] Willard Van Orman Quine. "Variables Explained Away." Proceedings of the American Philosophical Society (1960). Reprinted in Selected Logic Papers, Ed. by Quine. New York: Random House, 1966: pp. 227-235.

[Quine 1972] ___________. "Algebraic Logic and Predicate Functors", in Logic and Art, Ed. by Rudner and Scheffler. Indianapolis: Bobbs Merrill, 1972: pp. 214-238. Condensed version in Proceedings of the Second Scandinavian Logic Symposium, Ed. by Fenstad. Amsterdam, North Holland Publishing Co., 1971: pp. 304-316.

[Rasiowa, Sikorski 1963] Helena Rasiowa and Roman Sikorski. The Mathematics of Metamathematics, Panstwowe Wyndawnictwo Naukowe, 1963.

[Reichenbach 1947] Hans Reichenbach. Elements of Symbolic Logic. London: Collier-MacMillan, 1947. New York: Free Press Paperback, 1966.

[Scharle 1975] Thomas Scharle. "Axiomatization of Fragments of S5". Notre Dame Journal of Formal Logic, 16, (1975).

[Schönfinkel 1924] Moses Schönfinkel. "Uber die Bausteine der Mathematischen Logic." Translation in From Frege to Godel, Ed. by Van Heijenoort. Cambrdige: Harvard University Press, 1967.

[Scott 1968] Dana Scott. "Addrice on Modal Logic." In Philosophical Problems in Logic, Ed. by K. Lambert. Dordrecht: D. Reidel Publishing Co., 1970: pp. 143-175.

[Scott 1971] _________. "On Engendering an Illusion of Understanding." Journal of Philosophy 68 (1971): pp. 787-807.

[Scott 1972] _________. "Background to Formalization." In Truth, Syntax and Modality, Ed. by Leblanc. Amsterdam North Holland Publishing Co., 1972: pp. 244-273.

[Scott 1974] _________. "Rules and Derived Rules." In Logical Theory and Semantic Analysis, Ed. by S. Stenlund. Dordrecht: D. Reidel Publishing Co., 1974: pp. 147-161.

[Scott 1975] Dana Scott. "Completeness and Axiomatizability in Many-Valued Logic." In Proceedings of Symposia in Pure Mathematics, Vol 25: Proceedings of the Tarski Symposium, Ed. by L. Henkin. Providence: American Mathematical Society, 1975: pp. 412-435.

[Segerberg 1971] Krister Segerberg. An Essay in Classical Modal Logic. Ph.D. Dissertation, Stanford University, 1971.

[Segerberg 1973] ___________. "Two Dimensional Modal Logic," Journal of Philosophical Logic, 2, (1973): pp. 77-96.

[Thomason 1974] S.K. Thomason. "An Incompleteness Theorem in Modal Logic." Theoria, 60, (1974): pp. 30-35.

[Vlach 1973] Frank Vlach. 'Now' and 'Then': A Study in the Logic of Tense Anaphora. Ph.D. Dissertation. U.C.L.A., 1973.

[Wang 1952] Hao Wang. "Logic of Many-Sorted Theories." The Journal of Symbolic Logic, 17, (1952): pp. 105-116.

[von Wright 1951] Georg Henrik von Wright. Essay in Modal Logic. Amsterdam: North Holland Publishing Co., 1951.